DOCUMENT RESUME

ED 475 202 UD 035 607

AUTHOR	Currie, Janet; Moretti, Enrico
TITLE	Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings and Longitudinal Data. NBER Working Paper Series.
INSTITUTION	National Bureau of Economic Research, Cambridge, MA.
REPORT NO	NBER-WP-9360
PUB DATE	2002-11-00
NOTE	66p.
AVAILABLE FROM	National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138. Web site: http://www.nber.org.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	*Access to Education; Child Health; *Educational Attainment; Health Behavior; Higher Education; *Human Capital; Infants; Marital Status; *Mothers; *Parent Influence; Prenatal Care; Smoking; Womens Education

ABSTRACT

This study estimates the effect of maternal education on birth outcomes using data from the Vital Statistics Natality files for 1970 to 1999. It also assesses the importance of four potential channels through which maternal education may improve birth outcomes: use of prenatal care, smoking behavior, marriage, and fertility. In an effort to account for unobserved characteristics of women that could induce spurious correlation, the study pursues two distinct empirical strategies. First, it constructs panel data by linking women in different years of the Vital Statistics records and examines the effects of changes in education on changes in birth outcomes. Second, it compiles a new data set on openings of two and four year colleges between 1940 and 1990. The study uses data about the availability of colleges in the woman's county in her 17th year as an instrument for maternal education. Findings using the two approaches are similar. Higher maternal education improves infant health, as measured by birthweight and gestational age. It also increases the probability that a new mother is unmarried, reduces fertility, increases use of prenatal care, and reduces smoking, suggesting that these are important pathways for the ultimate effect on health. (Contains 43 references.) (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 475 202

NBER WORKING PAPER SERIES

MOTHER'S EDUCATION AND THE INTERGENERATIONAL TRANSMISSION OF HUMAN CAPITAL: EVIDENCE FROM COLLEGE OPENINGS AND LONGITUDINAL DATA

Janet Currie
Enrico Moretti

Working Paper 9360
http://www.nber.org/papers/w9360

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2002

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. Currie
National Bureau of Economic Research

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Tom Kane, Claudia Goldin, Lance Lochner, Justin McCrary and seminar participants at UCLA, Berkeley, London School of Economics, IZA, Heidelberg University, NBER Summer Institute, Columbia, CUNY, UC Davis, University of Washington at Seattle, the University of Colorado at Denver, UC San Diego, Princeton and Yale provided very helpful comments. We also thank Michael Greenstone and Ken Chay for sharing their Vital Statistics data and David Card for sharing his data on colleges that closed. Rebecca Acosta, Ben Bolitzer and Anna-Maria Bjornsdotter provided excellent research assistance. We thank NIH (Currie) and the UCLA Senate (Moretti) for financial support. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2002 by Janet Currie and Enrico Moretti. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

UD 035 607

BEST COPY AVAILABLE

Mother's Education and the Intergenerational Transmission of Human Capital:
Evidence from College Openings and Longitudinal Data
Janet Currie and Enrico Moretti
NBER Working Paper No. 9360
November 2002
JEL No. I1, I2, J0

ABSTRACT

We estimate the effect of maternal education on birth outcomes using data from the Vital Statistics Natality files for 1970 to 1999. We also assess the importance of four potential channels through which maternal education may improve birth outcomes: use of prenatal care, smoking behavior, marriage, and fertility. In an effort to account for unobserved characteristics of women that could induce spurious correlation, we pursue two distinct empirical strategies. First, we construct panel data by linking women in different years of the Vital Statistics records and examine the effects of changes in education on changes in birth outcomes. Second, we have compiled a new data set on openings of two and four year colleges between 1940 and 1990. We use data about the availability of colleges in the woman's county in her 17th year as an instrument for maternal education

Our findings using the two approaches are similar. Higher maternal education improves infant health, as measured by birthweight and gestational age. It also increases the probability that a new mother is married, reduces parity, increases use of prenatal care, and reduces smoking, suggesting that these are important pathways for the ultimate effect on health.

Janet Currie
Department of Economics, UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1477
and NBER
currie@simba.sscnet.ucla.edu

Enrico Moretti
Department of Economics, UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1477
and NBER
Moretti@econ.ucla.edu

Introduction

Over the past half century, American women have experienced tremendous increases in average educational attainment. Returns to education are generally quantified in terms of increases in wages. However, in addition to earning more, well educated women marry better, and may use any given set of health inputs more efficiently, by for example, reducing unhealthy behaviors such as smoking. If higher maternal education does indeed improve child health outcomes, then conventional estimates of the returns to education, which focus only on wages, understate the social benefits. Moreover, to the extent that healthier children go on to be more productive and more highly educated adults themselves, there will be an important inter-generational spillover that analyses of wage effects alone will not capture.[1] On the other hand, college educated women tend to have fewer children (and sometimes none at all), and some researchers have suggested that by increasing the opportunity cost of a woman's time, gains in women's education above some level might negatively affect some measures of child well-being. Thus, the extent to which gains in maternal education have translated into improved child quality is an open question.

This paper makes two contributions. We estimate the effect of maternal education on birth outcomes using data from individual birth certificates from the Vital Statistics Natality files for 1970 to 1999 and find that infants of college educated women have better outcomes, measured in terms of both birth weight and the incidence of prematurity. We also assess the importance of four potential channels through which maternal education may improve birth outcomes. First, we examine the effect of education on the use of prenatal care. Because more educated women earn more, they may be able to afford more or better health care. Second, more educated women are also likely to marry higher earning men, which will further raise family income. Hence, we examine the extent to which increase in a woman's education raises the probability of marrying a highly educated man. Third, education may induce women to have healthier behaviors. In particular, we analyze the effect of education on the probability of smoking during pregnancy and find large effects. Finally, we look at

[1] Lleras-Muney (2001) makes this point in a somewhat different context, showing that there is a return to education in the form of longer life-expectancy among adults. Lazear (1983) provides a theoretical analysis of intergenerational externalities in education.

the effect of education on fertility. As Becker's quality/quantity trade-off suggests, education may induce women to have fewer children of higher quality.

However, it is not obvious that these relationships are causal since better educated women differ in many ways from less educated ones. In an effort to isolate the causal effect of education, we pursue two distinct empirical strategies. First, we construct panel data by linking women in different years of the Vital Statistics and examine the effects of *changes* in education on *changes* in birth outcomes. These longitudinal models control for permanent unobserved factors such as genetic endowment, family background, ability, and discount rates. However, they are not necessarily robust to the presence of transitory shocks that may be correlated with the decision to go back to school after the first birth.

Second, we turn to an instrumental variables strategy. Following the literature that uses institutional features of the college market as instruments for educational attainment, we have compiled a new data set on openings of two and four year colleges between 1940 and 1990. We use this data about the availability of colleges at the county level as an instrument for maternal education in models of birth outcomes and in models of potential mediating factors such as prenatal care, smoking behavior, marriage, and fertility. Our instrumental variables (IV) models control for many potentially unobserved confounding factors by including county-year-of-birth fixed effects, so that our estimates are identified by differences in the availability of educational services among different cohorts of women delivering in the same county and year.

One could argue that colleges tend to open in counties where residents' education is already increasing or is expected to increase, and therefore are not a cause but an effect of increasing education. Although we cannot completely rule out this possibility, we provide seven pieces of evidence which suggest that college openings do in fact increase educational attainments. First, we find that maternal education and birth outcomes have no clear trend in the ten years before the opening of a college, but improve afterwards. Second, while the opening of a college in the mother's 17^{th} year increases her educational attainment, openings in her 25^{th} or later years do not. Third, the opening of four year colleges increases four year graduation rates, but does not affect the probability of having some college, while the opening of two year colleges increases the probability of having some college but has a much smaller effect on four-year graduation rates. Fourth, when male-only

colleges become coed, maternal education increases, but when female-only colleges become coed, we find no effect on women's education. Fifth, the opening of public colleges has a larger effect than the opening of private colleges, and sixth college openings matter more in relatively "college poor" counties. Finally, controlling for the mean education of father's in the county when the woman was 17 as an additional index of demand for education has little effect on our estimates.

Although the instrumental variable and longitudinal model estimates are based on completely different samples and identifying assumptions, our findings are remarkably similar using both methods. Higher maternal education improves child quality, as measured by birth weight and gestational age. It also increases the probability that a new mother is married, is associated with higher husband education, reduces parity, increases use of prenatal care, and reduces smoking, suggesting that these are all important pathways for the ultimate effect on health. As one might expect, the estimated effects are larger for less educated women, but remain substantial even among college graduates.

The rest of the paper is laid out as follows. Section 2 provides essential background, while Section 3 lays out a simple model showing the inter-relationships between the outcomes we examine. Section 4 discusses the data and Section 5 describes our empirical strategy. Results appear in Section 6 and 7, which are followed by the conclusion.

2. Background

Our work builds on a large literature examining the role of maternal education in the "production" of child quality, as well as the effect of female education on marriage and fertility. In addition, to the extent that better child health is viewed as a "return" to maternal education, our work is in the tradition of research measuring the returns to college using changes in the accessibility of college as an instrument. This section briefly reviews key findings from these literatures.

a) Maternal Education and Child Quality

Many studies report a correlation between maternal education and measures of child health. In fact, this robust relationship was one factor underlying the World Bank's drive over the past decade to promote maternal education in developing countries (World Bank, 1993). Much of the

existing literature differs from this paper in that it focuses on developing countries, emphasizes the effects of improvements in relatively low levels of education, and does not attempt to establish whether the effect of maternal education is causal.[2]

There is little work exploring the relationship between maternal education and child health in developed countries. There is however, some work exploring the effects of maternal education on children's test scores and schooling attainment in the U.S. Rosenzweig and Wolpin (1994) use a sample of siblings to examine the effect of maternal schooling on children's achievement test scores. They find that mothers who continued their schooling between births increased the test scores of the younger siblings without harming the older ones. We will pursue a similar strategy using the panel data set we have constructed.

On the other hand, Behrman and Rosenzweig (2002) conclude that much of the positive cross-sectional relationship between the schooling of mothers and children is due to heritable ability as well as assortative mating. They note that increased schooling of mothers reduces the amount of time mothers spend in the home (i.e. time inputs into child production) which could well have negative effects on child outcomes. In keeping with this hypothesis, studies that have directly examined the effect of maternal employment on measures of child well-being often find some evidence of negative effects, though the results are inconsistent across gender and age groups and/or outcome measures, and it is difficult to control for the endogeneity of maternal employment.[3]

If maternal education does affect child health, then this may have significant intergenerational effects. Poor child health affects adult health, and may also have indirect effects through reductions in cognitive functioning and/or schooling. There are many studies linking low birth weight to future cognitive deficits, for example (see Currie, 2000 for a review of some of this literature). Currie and Hyson (1999) show using data from the 1958 British cohort study that low birthweight also has long-term negative impacts on test scores, employment probabilities, and wages among young adults.

Thus, the existing literature suggests that increases in even high levels of maternal education

---

[2]Two papers that do address the question of causality are Desai and Alva (1998) and Thomas, Strauss and Henriques (1991). The first suggests that much of the return to female education may be through higher father's earnings, while the second finds that female education is associated with factors such as reading newspapers, and thus seems to have a direct effect on the ability to gather information about health.

[3]See Blau and Grossberg, 1992; Desai, Chase-Lansdale, and Michael, 1989; Neidell, 2000; Parcel and Menaghan, 1994; Ruhm, 2000; Waldfogel, Han, and Brooks-Gunn, 2002.

might have a positive effect on infant health by improving maternal behaviors that are linked to poor infant health outcomes. On the other hand, increases in the value of maternal time with education may militate against time-consuming investments in infant health. The literature also highlights the dangers in interpreting any observed correlations between maternal education and child health inputs or outputs as causal. Hence, the relationship between increases in college attendance among American women and infant health outcomes remains an open question.

b) Education, Assortative Mating, and Fertility

Education may affect child outcomes either directly or indirectly through its effects on the choice of father. Becker (1981, 1985) predicted that more educated men would be likely to marry less educated women because of gains to specialization in "household production".[4] However, the empirical evidence suggests the opposite: Highly educated men and women tend to marry each other. Mare (1991) examines data from five decades and concludes that the tendency towards "positive assortative mating" became stronger between the 1930s and the 1970s, and then remained fairly constant during the 1980s. Much of the effect seems to be due to the fact that college educated people find potential spouses among their college classmates.

Goldin (1992, 1995) examines the implications of positive assortative mating on a woman's "returns" to higher education. She finds that for women, much of the return to college may come in the form of a higher earning spouse. For the cohort that graduated between 1945 and 1960, roughly half of the gain to household income was through this channel. Moreover, of women who pursued a career subsequent to college, some remained childless, suggesting that in our sample of women bearing children, a large part of the return to college was through enhancements in spouses' earnings rather than increases in their own earnings.[5]

Education may also affect child outcomes indirectly through its effects on fertility. Becker and Lewis's (1973) discussion of tradeoffs between the "quantity" and "quality" of children desired is often given as an explanation of the negative relationship between women's education and their

---

[4] Since Becker, several authors have developed theoretical explanations for positive assortative mating. See Lam (1988) and Oppenheimer (1988).

[5] Were we to use these data to study maternal health, selection induced by education would be a significant problem. However, we have the whole universe of infants born. Even if college education causes some women not to bear children, it may still have a positive effect on the health of those infants who are born.

fertility. The model also implies that increases in education will be associated with improvements in child quality.

In summary, there is surprisingly little evidence available about whether increases in maternal education in a country like the United States have a causal effect on infant health outcomes. The literature suggests that in addition to any direct effects of education on child health production, it is important to consider indirect effects through assortative mating and fertility decisions.

c) College Availability and Schooling

Given that we examine both two and four-year college openings, studies by Kane and Rouse (1995) and Rouse (1995) are particularly relevant. Kane and Rouse were the first to demonstrate that there was indeed a labor market return to attendance at two-year colleges. Rouse shows that access to two year colleges increases overall educational attainment, though it does not affect the probability of attaining a four year degree. She also finds that increasing access to two-year colleges tends to divert some students from four-year to two-year colleges. Our results on the effects of two-year relative to four-year colleges on women's educational attainment are consistent with Rouse's findings.

Card (1995) uses college proximity as an instrumental variable to estimate returns to schooling. He shows that instrumental variables estimates are larger than OLS estimates, which is due to there being high returns to education among those most constrained by lack of college availability. We return to this point below.

Finally, Goldin and Katz (1999) document the causes and effects of the growth in the availability of public universities in the period 1890-1940, which immediately precedes the period that we focus on in this paper (we consider college opening that occur between 1940 and 1999). Consistent with what we find in the later period, Goldin and Katz find that the great expansion in the number and size of state funded public universities that took place in certain states between 1890 and 1940 significantly increased access to higher education and college enrollment. They also argue that in the first half of the century, the high school movement brought students from less privileged backgrounds to college and thus further buoyed enrollments in the public universities.

3. A Simple Model of the Effects of Maternal Schooling on Family Structure, Health Inputs, and Child Health Outcomes.

The model in this section shows the relationships between the different outcomes we will examine. It is based on the idea that women "produce" child health outcomes by combining inputs in ways that are circumscribed by the available production technology.[6] Women are assumed to derive utility from child "quality", as well as from their own consumption and leisure:

(1) $U = U(Q, C, L; X, u)$,

where Q is the stock of child quality, C is consumption of other goods, L is leisure, X is a vector of individual characteristics, and u is a vector of neighborhood characteristics that may affect tastes such as average family size in the area. Women's utility is increasing in all three arguments. $Q = f(N,q)$ where N is the number of children and q is the quality per child. It is not important that any particular functional form be assumed. Rather, as Becker and Lewis (1973) point out, what is important is that "an increase in quality is more expensive if there are more children because the increase has to apply to more units; similarly, an increase in quantity is more expensive if the children are of higher quality, because higher-quality children cost more" (page S280).

Utility is maximized subject to the following set of constraints:

(2) $q = q(G, V; S, X, v, e)$,

where G represents goods invested in child quality, V represents time invested in producing child quality, S is the mother's schooling, v is a vector of neighborhood characteristics such as medical services that may affect child quality, and e is a random shock. Equation (2) can be thought of as a production function for child quality. Education is expected to improve the productivity of mothers. That is, the more educated mother is assumed to be able to produce a higher "quality" child with a given budget of goods and time.

Mothers also face the following set of constraints:

(3) $C = Y - P^{*}G$,

---

[6] The model derives from Grossman's (1970, 1999) model of health as human capital. For early examples of its application to infant health see Rosenzweig and Schultz (1982, 1983, 1988).

where P are exogenously determined prices,

(4) $Y = Y_f + w \cdot H$,

where Y_f is the amount of income that the woman receives from the father of her child,

(5) $w = w(S; z)$,

where z is a vector of local labor market conditions, $w'(S)>0$, and

(6) $L + N^{*}V + H + M = 1$,

where M is the time cost of being married (if a woman is married) and the amount of time available is arbitrarily normalized to one. If a woman is married, then $M>0$ and $Y_f=Y_f(S)$ where $Y'(S)>0$. That is, we assume that a woman can find a better quality husband if she is more educated. If she is not married, then $M=0$ and $Y_f=0$. In order to solve this problem, the woman must determine whether her maximum utility will be higher in the married or unmarried state.

The mother's choice variables are S, M, N, G, V, and H. We will observe analogues of all of these measures except H in our data. S and M are directly observed, and while we do not know the completed number of children, we can observe the birth order of each child in our data to see whether, on average, less educated women are having births of higher parity. G and V are proxied by the whether or not prenatal care began in the first trimester of pregnancy, which can be viewed as an indicator of the mother's willingness to invest in the unborn child and also by whether or not the mother smokes.

It is interesting to compare the effects of these two health inputs. Prenatal care is costly so that increases in prenatal care with education might reflect a relaxation of the budget constraint that occurs with increased wealth. However, refraining from smoking saves money and time so that reductions in smoking are unlikely to be accounted for by increases in wealth alone, and may reflect an effect of education via the production function. Moreover, smoking is the leading cause of low birthweight in the United States, so it is an interesting outcome to examine in its own right: Lightwood, Phibbs, and Glantz (1999) estimate that up to a quarter of low birthweight infants born in the 1980s were adversely affected by maternal smoking while pregnant.

We can now consider the effect of reducing the cost of obtaining a college education on each of these outcomes. Other things being equal, a reduction in cost should increase the quantity of education, S, consumed. The increase in schooling increases the expected wage, increasing the

opportunity cost of time spent in non-work activities such as child rearing and marriage-building. On the other hand, schooling increases the payoff to investing in marriage by increasing the value of the spouse that can be attracted. These considerations suggest that the effect of education on marriage rates, M, is theoretically ambiguous while the effect on the total amount of time devoted to child-raising, N*V, is negative.[7]

The model highlights the many pathways through which increased maternal schooling can affect our measures of child quality, q, which are birthweight and gestational age. Increased schooling will influence the probability of marriage, family income and family size, which will in turn influence the choice of inputs into child development. In addition, education may change the production function, allowing mothers to combine a given set of time and money inputs into outputs more efficiently.

The model sketched above can be solved to yield conditional demand functions for our indicators of M, N, G, V, and for our measures of child quality, q, of the following form:

(7) M, N, G, V, q = f(S, X, P, u, v, z, e)

where the demands/outputs are conditional on the endogenously chosen level of education, S, which in turn is assumed to depend on the availability of college. In most of our work, we will proxy the neighborhood-specific variables P, u, v and z using county-year fixed effects and mother-cohort trends. In the next section, we discuss the application of this model to our data.

4. Data

Our primary source of data is individual-level Vital Statistics Natality records from 1970 to 1999. These data are drawn from birth certificates and are thought to cover virtually all births in the United States. Most states currently report the mother's education, information about when she began prenatal care during pregnancy, whether she smoked during pregnancy, marital status, information about the fathers' education, and information about birth outcomes including

[7]Child-raising time can be reduced either by having fewer children, or by devoting less time to each child, or both. However, decreases in V must be offset with increases in G if a given level of q is to be maintained. Of course, to the extent that better educated women are able to combine G and V more efficiently in the production of q, quality may rise even with G and V constant.

birthweight and gestational age. One limitation of the Vital Statistics data is that the number of states reporting some types of information has increased over time, and thus we do not have a balanced panel of counties over time, and the number of counties is not the same for all our outcome measures. In the Data Appendix we describe which of the variables are not available in all years. The effect of this selection on our estimates is mitigated because we include either mother fixed effects or county*year dummies in all of our models.

Because there are well known differences in average outcomes by birth order, we focus most of our analyses either on first births (in the OLS and IV) or on changes between first and second births (in the panel). (The sole exception is when we examine parity, where we include births of all parities in the sample). We focus on white mothers, since our identification strategy does not work well for black mothers.[8]

In this paper, we use two separate identification strategies: longitudinal models and instrumental variables. The samples used in the two strategies are different, although they are both drawn from the Vital Statistics files. We now describe the two samples.

a) Longitudinal Sample

Our first set of analyses is based on a panel data set constructed from Vital Statistics records. The Vital Statistics system does not link mother's records between births, and we are not aware of previous attempts to create a panel from these data. However, it is possible to link first and second births with a fair degree of certainty for a subset of mothers using information about the mother's county of residence, the mother's year and state of birth, as well as each child's date of birth. More specifically, the longitudinal sample was obtained as follows: first, women with parity equal to 1 were selected. Then, women with parity equal to 2 were matched to women with parity equal to1 based on county, date of birth of first child (month and year, which was reported only up until 1993), mother's age at first birth, mother's year of birth and mother's state of birth.[9] Finally, only cells with a unique observation in both years were kept (one-to-one matches). The sample is therefore a

---

[8]Black women do not appear to be as strongly affected by new college openings as white women, perhaps because of segregation in the first part of the period under consideration. We hope to explore the relationship between education and birth outcomes among black women in future work.

[9]Because we are looking at white mothers only, we are implicitly using race for the match. Mothers at second birth are asked the year and the month of their previous delivery. The information on the month of birth of the previous delivery was dropped after 1993 as a result of budget cuts.

panel of mothers with two children, observed at first and second birth. For this analysis, we use all women aged 16 to 24 at first birth and 17 to 30 at second birth, since we want to capture all women who experienced a change in educational attainments between the births. Because we are including only one-to-one matches, mothers from smaller counties are more likely to be in the sample.[10] See the Data Appendix for further details.

Summary statistics for all first births in the Vital Statistics sample are shown in column (1) of Table 1. Although we don't use this sample in our analysis, column (1) is reported for the sake of comparison with the panel data and IV samples. Column (2) shows first births from the panel of first and second births that we are able to match. The matched longitudinal sample is not perfectly representative of the population. The differences arise because the women we can match one-to-one are more likely to be from rural areas than the full sample of women, and because mothers who change county between births cannot be matched.[11] Column (3) shows means for the second births in our panel. On average, women in the panel gained a half year of education between the first and second birth. There is a striking reduction in the probability of low birthweight or prematurity between first and second births. Table 1 also shows that the probability of using prenatal care in the first trimester and of being married increased between first and second births. [12]

The increases in schooling are explored further in Appendix Table A1, which shows the distribution of changes in years of schooling between first and second birth. The table indicates that 14.5% of the mothers add one year, 5.9% add 2 years, and 1.6% add 3 years. One concern is that these gains may in part reflect measurement error. Specifically, there are two potential sources of measurement error. First, as in other datasets, there may be misreporting of schooling by mothers. This type of measurement error will introduce the standard attenuation bias, leading us to underestimate the true effect of schooling on outcomes. Second, measurement error may arise from imperfect matching of mothers across births: We may erroneously believe that we have matched the

---

[10] We have re-estimated all longitudinal models using a sample that includes *all* cells that can be matched, not only cells with size equal to one. Results did not change.

[11] The comparison of column (1) and (2) indicates that the mothers we can match are less educated than the full sample, less likely to have low birth weight or pre-term babies, less likely to get prenatal care in the first trimester and more likely to be married.

[12] In a previous version, we also examined the probability that infants were low birthweight conditional on being full-term, since the effect of inter-uterine growth retardation is different than that of preterm delivery, although both may result in low birthweight. The results were similar to those we report for low birth weight.

same mother across births, while in fact we have matched two separate individuals. This second source of measurement error is more worrisome, because it biases our estimates toward the OLS estimates.

To gain a sense of how large this problem is, we compare the changes in schooling in our sample with those in the National Longitudinal Survey of Youth (NLSY). To maximize comparability, we take all white mothers in the NLSY who do not change county of residence between births; and in our Vital Statistics panel, we take all mothers born between 1957 and 1964. Appendix Table A2 indicates that the distribution of educational changes in two samples is quite similar, although on average our panel reports slightly more changes than the NLSY. About 38 percent of mothers in our sample who are 18 or younger at first birth report a positive change in schooling, compared to 36 percent in the NLSY. For mothers who were 19 or 20 at the first birth, the corresponding figures are 17 percent and 15 percent respectively. In the older group of mothers, our sample overestimates the probability of increases in schooling relatively to the NLSY. But because most changes in schooling happen at earlier ages, most of the identification of our longitudinal models comes from the younger groups. (Estimates for the subset of mothers 21 or younger are similar to estimates presented here). We conclude that our longitudinal sample contains some measurement error of the second type (mismatching), but probably not enough to drive our results. Below we show that measurement error of the first type (misreporting of schooling) is no worse than it is in the Census of Population.

b1) IV Sample

We now turn to the sample used to estimate the IV models. While in the longitudinal sample we need mothers of school age (since identification comes from mothers whose education changed between the first and second birth), in our instrumental variables strategy we focus on first-time mothers 24 to 45 years old in order to select women whose education is likely to have been completed. We take a 10% random sample of the full sample of women in this category for ease of estimation.[13] Column (4) of Table 1 shows means for this sample. Compared to the full sample of

[13]Since our instrument varies at the county-cohort of birth level, the effective sample size for the instrumental variable estimator is the number of cohorts times the number of counties. Using only 10% of all births greatly reduces computation time but has no effect on the effective sample size.

first-time mothers, these older mothers are more educated, less likely to have negative outcomes, more likely to get prenatal care in the first trimester, much more likely to be married, and less likely to smoke. They also have husbands who are more educated on average. However, the median income in their county in the year when they are 17 is more similar to the full sample of mothers than to that of the mothers in the panel, and the number of two and four year colleges is also comparable to the full sample.

The last four columns break out this sample of older mothers by education. In all cases, more education is associated with more desirable birth outcomes. The monotonic relationship between favorable birth outcomes and maternal education is very striking. These improvements in outcomes may come about through many channels, as illustrated by the monotonic relationship between education and early prenatal care, probability of smoking during pregnancy, probability of being married, husband education and fertility. Just to take one example, only 2 percent of college educated mothers in this age group smoke during pregnancy. The corresponding figures for mothers with some college, high school and less than high school are 8 percent, 17 percent, and 34 percent. College educated women are also more likely to come from counties with a higher per capita number of four-year colleges in the year when they were 17.

In interpreting the IV estimates, it is important to realize that the IV sample is not representative of the entire population of mothers. Very few high school drop-outs have their first baby at 24 or later, while many college graduates do. As we show below, we have experimented with a sample that includes mothers 22 and older and have found similar results. The advantage of the sample that includes mothers 22 to 45 over the one that includes mothers 24 to 45 is that many women have their first baby between 22 and 24, so that this larger sample is significantly more representative. The disadvantage (and the reason why we don't use this sample as our main sample) is that many women are still in college at 22. It is worth noting that the fact that our IV sample is not representative does not affect the internal validity of the IV estimates. Moreover, results based on the IV sample of older mothers are complemented by those based on the panel, which includes younger mothers.

While we have no reason to believe that education is systematically misreported in the Vital Statistics data, we have compared education in the Vital Statistics to Census reports. Appendix

Table A3 indicates that the education distribution in the Vital Statistics data is consistent with the education distribution of first time mothers in the Census. See the Data Appendix for a more detailed discussion of this Table.

b2) Data on College Openings

Our IV analysis combines Vital Statistics natality data with a unique data set on the number of two and four-year colleges by county and year that we compiled for this research. The construction of this data set is described in the Data Appendix. Figure 1 shows the tremendous rise in the number of colleges between 1940 and 1996. In 1940, there were 346 two-year colleges in our sample and 1301 four-year colleges. By 1996, these figures had risen to 1436 and 1808 two and four-year colleges respectively. The pattern for two-year colleges is more "S-shaped" than that for four-year colleges reflecting the great expansion of the two-year college system during the 1960s and 70s.[14]

We use these data to construct measures of the availability of two and four-year colleges. Each measure is the number of colleges which existed in the woman's county when she was 17 years old, divided by the estimated number of 18 to 22 year olds in the county in that year (in thousands). This instrument takes into account the fact that cohort size is likely to have an impact on the availability of college given any fixed number of schools (c.f. Card and Lemieux, 2001; Welch, 1979).[15] We construct separate measures for two and four-year colleges.

The ideal instrumental variable would account not only for the number of colleges, but also for their size. Information on the availability of classroom space or capital expenditures is not available for part of the period under consideration. Information on enrollment is available, but enrollment reflects both the supply of college places and the demand for these places, and thus is not a valid instrument. But it should be noted that this is more an issue of power of the instrument than of validity of the instrument. Because our first stage using the number of colleges is fairly precisely estimated, our instrument, although not perfect, seems adequate in this respect.

The most important limitation of the Vital Statistics data is that we observe the mother's

---

[14]Virtually all of the two-year colleges in our sample are public institutions, while roughly half of four-year colleges are public.

[15]The number of individuals 18 to 22 years old is from the Census. See the Data Appendix. Note that measurement error in our denominator should not carry over to our left hand side variable since they are measured in 2 different sources.

residence at the time of the birth of the baby, rather than at the time when she was 17. Thus, we are forced to assume that her county of residence is the same at the birth as it was at 17. Because young women have high mobility rates, this assumption is potentially problematic. If mothers randomly change location between age 17 and the time of the birth, then we will tend to understate the extent to which college openings affect women's educational attainment in the first stage regressions. However, 17 year old women may move in order to attend college and then stay in the new location, or counties that experience college openings may be become more attractive to college educated women.

We cannot completely rule out these two possibilities. But we present two pieces of evidence that suggest that this type of mobility is not driving our results. First, if college educated women tended to move to counties with college openings (for example, because of employment opportunities), then we would see a correlation between college openings and college availability measured at ages greater than 17. We test for this possibility by examining the relationship between maternal education and college availability at 25, 30, 35 and 40. We find no evidence of a positive relationship between maternal education and college availability at older ages.

It is still possible that our results are driven by the migration of 17 or 18 year old women seeking education to counties with recent college openings. While highly selective institutions recruit students from across the country, most of the four year and virtually all of the two year colleges in our sample are not of this type. A typical new institution is more likely to resemble Coleman College (La Mesa, CA) or Aims Community College (Greeley, CO) than Harvard or Princeton. Although most of the schools in our sample are non-selective institutions that one would expect to enroll mainly local women, it is still possible that at least some of our results are explained by the endogenous mobility of college students.

While we cannot completely rule out this possibility, we use Census data in an attempt to quantify the magnitude of the bias introduced by endogenous migration. The Census has the advantage that it provides information on location of residence five years prior the Census year. We use such information to try to replicate the first stage based on location of residence at age 14 to 16, i.e. before women move to go to college. We find that estimates for four-year colleges based on the number of colleges that exist in the location of residence at age 14 to16 are similar to estimates based

on the number of colleges that exist in the current location of residence. In interpreting these findings, one should keep in mind that Census data are far from ideal, because they are available only for a limited number of years, and identify "PUMAs", not counties. See the Data Appendix for details.[16]

5. Identification Strategies

In this paper we adopt two separate identification strategies to estimate the effect of schooling on child quality, and the channels through which schooling affect child quality. Although neither identification strategy is based on incontrovertible assumptions, the two strategies depend on very different assumptions. Therefore, we hope to shed some light on the validity of our identification assumptions by comparing estimates from the two strategies.

a) Longitudinal Models

Using panel data, it is straightforward to estimate models that control for *permanent* unobserved characteristics of mothers by estimating equations of the following form:

(1) $d\text{OUTCOME} = a_0 + a_1 * d\text{EDUC} + a_2\text{AGE}_1 + a_3\text{AGE}_2 + a_4\text{COHORT} + v$,

where dOUTCOME is the change in one of our outcome measures, dEDUC is the change in the mother's years of education between the births, AGE_1 is a vector of dummy variables for the mother's age at the time of the first birth, AGE_2 is a vector of dummy variables for the mother's age at the time of the second birth, COHORT is a vector of indicators for the mother's decade of birth (e.g. 1920-29, 1930-39, etc.). It is important to control for age, since the risk of poor infant outcomes as well as marriage and fertility decisions depend on age (very young mothers are less likely to be married, and both very young and very old mothers are more likely to have poor outcomes). The specification of the age variables in equation 1 allows the effect of birth interval to

[16]One final caveat is needed. In the panel, we include women who have two births, who are more likely to be married than women in the full sample of 16-24 year old women (column 1). Similarly, in the IV sample, we include women who are 24 or older who, again, are more likely to be married. Therefore our results on marriage may be less generalizable than some of the other results.

differ by the mother's age at first birth. While cohort dummies would fall out of a straight fixed-effects model, model (1) allows the effects of changes in education to differ for different cohorts. The cohort of the mother proxies for many factors such as attitudes towards education, marriage, female employment, and child bearing. We obtained very similar estimates with a specification which also includes county dummies.

The model in equation (1) is identified by comparing mothers whose education does not change with mothers whose education increases between the first and second birth. The model is based on the assumption that the only source of heterogeneity that affects outcomes is *permanent*. In particular, model (1) controls for differences in ability, genetic endowment, and family background. However, it is possible to think of cases where this assumption fails. The presence of *transitory* shocks that affect both education and birth outcomes could be problematic. For example, women who decide to go back to school after having the first child may be experiencing improvements in their economic conditions relative to women who decide not to go back to school. In this case, model (1) would overestimate the effect of schooling. For this reason we now turn to a different identification strategy, based on instrumental variables. Although the IV estimator is based on a different set of assumption, we will see that IV estimates are generally consistent with estimates obtained from longitudinal models.

b) Instrumental Variables Models

The first stage in our instrumental variables regressions take the form:

(2) EDUC = b_0 + b_1IV-2 + b_2IV-4 + b_3AGE + b_4COHORT + b_5COUNTY*YEARBRTH + b_6 INCOME17 + b_7 URBAN17 + u,

where IV-2 and IV-4 denote our instrumental variables for two and four-year colleges; AGE is a vector of single year of maternal age dummies as in (1); COUNTY*YEARBRTH is a vector of indicators for the county and year of the child's birth; INCOME17 and URBAN17 control for the median income and the percent urban in the county when the woman was 17, and u is a random error term.

The county/year effects control for many characteristics of the local area that may affect outcomes, such as the availability and quality of medical services, the local business cycle, pollution,

etc. Identification comes from the fact that *within each county and year of birth of the baby*, there are mothers who were 17 before a college opening, and mothers who were 17 after a college opening.

The income and urbanicity measures control for some of the factors likely to affect the educational attainment of the mother's cohort. In some of the models, we test the robustness of our results by including state-mother cohort trends, which allow the cohort effects to vary from state to state, and by including the average education of fathers in the county (of all ages) when the mother was 17 as a measure of demand for education (as discussed further below).

Our second stage regression models take the following form:

(3) OUTCOME = c_0 + c_1PEDUC + c_2AGE + c_3COHORT +c_4COUNTY*YEARBRTH + c_5 INCOME17 + c_6 URBAN17+ z,

where PEDUC is the predicted years of education from (2), and z is a random error. In all the models, the standard errors allow for potential correlations between the errors within county-year at 17 clusters.

This specification does not capture the part of the return to schooling that arises through residential location. More educated women can probably afford to live in areas with better hospitals and better environmental quality. By including county*year effects, we absorb the effect of these geographical factors on birth outcomes, and in this sense, we underestimate the benefit of schooling.

c) Validity of the Instrument and its Limitations

In contrast to most studies using policy changes as an instrument, the possible correlation between changes in number of colleges and *contemporaneous* changes in other factors affecting infant outcomes is not problematic. For example, if counties that experience increases in the number of colleges also experience increases in the number of hospitals, our IV estimates are still valid, because we control for county*year-of-birth effects. The instrument is not valid if changes in the number of colleges when mothers are 17 are correlated with unobserved factors that affect child quality years later, when mothers are 24 or older. The two main concerns regarding the validity of our instrumental variables are the fact that the geographical location of new colleges may not be random, and endogenous mobility of women who move to attend college. We now discuss these two

concerns.

It is possible that college openings reflect expected increases in demand for education. In this case, educational attainment would drive college openings, not vice-versa, and our estimates could overestimate the effect of schooling on birth outcomes. On the other hand, it is also possible that state legislatures place colleges in a compensatory way, i.e. in under-served areas that have low college attendance. For example, legislation establishing the new University of California Merced campus begins: "The San Joaquin Valley is the most populous region of the state without a University of California campus, and has one of the lowest rates of college participation of all regions in California. Access to postsecondary education is determined, in significant measure, by a student's proximity to college campuses" (California State Education Code, Section 92160). Similarly, other state plans for higher education emphasize increasing access to education as a main goal of public investments.[17] These policy statements suggest that to the extent that there is any correlation, the location of new public colleges may be negatively rather than positively correlated with average academic attainments in the area, which would lead us to underestimate the effect of education on birth outcomes.

The location of private colleges may be governed by different considerations. On the one hand, it is possible that private colleges are located where demand is anticipated to grow. On the other hand, private college location could be affected by land prices, which would dictate location in less developed areas. Ultimately, whether new colleges locate in areas with larger expected increases in schooling is an empirical question.

We assess the validity of our instrument with several specification checks. First, we look at whether education appears to be increasing before college openings. If openings determine education rather than vice-versa, we should find little evidence of a pre-trend in schooling in the ten years before the opening of a new college. We also look for the presence of pre-trends in birth outcomes.

Second, we ask whether new four-year colleges have a stronger effect on the completion of four-year college degrees than on the completion of two-year degrees, and whether new two-year colleges have a stronger effect on the completion of two-year degrees than on the completion of four-year degrees, as argued by Rouse (1995). For example, finding that a new four-year college has the same effect on the completion of four-year degrees and two-year degrees would be a sign of spurious

correlation, because it would suggest that new colleges open in counties where education is increasing at every level.

A third test of the validity of our instruments is to estimate first-stage models that include measures of college availability when the mother was 25 or older in addition to measures of college availability when the woman was 17. If our identification strategy is valid, then measures of opportunities at age 17 should have a larger effect on the woman's educational attainment than measures of college availability taken when she is past the usual age of college attendance. Conversely, if college educated women 25 or older move to counties that experience college openings, then measures of college availability taken at 25 or older should be stronger predictors of a woman's educational attainment than our measure of college availability measured at age 17.

Fourth, we ask whether the opening of new public colleges has a different impact than the opening of new private colleges. Looking at differences between private and public colleges is interesting for two reasons. First, if our story is correct, and our instrument captures reductions in the cost of attendance for the marginal woman, we expect the opening of a public college to have a larger impact than the opening of a private college. Women who decide to attend college just because a new college opens in their county are probably from disadvantaged backgrounds, and therefore are more likely to attend public colleges. Presumably, most women who attend private colleges would have gone to college anyway, even if there were no college in their county of residence. Finding larger effect for private colleges would cast doubt on the validity of our instrument.

Fifth, we turn to single-sex colleges. Ideally, we would like to use the opening of single sex colleges as a specification check. But there are very few openings of single sex colleges in the period under consideration. We focus on transformations of single-sex colleges into coed colleges. If our story is correct, male-only colleges that become coed increase women's access to college and hence should have a positive effect on female education; but female-only colleges that become coed should have no positive effect on women's education.

Sixth, we ask whether the opening of a new college has a greater effect on educational attainments in regions that are "college poor". That is, we expect that the opening of a new college should have a smaller effect if there are many colleges close by than if there are few. Seventh, we

---

[17] See examples on www.sheeo.org/govern/gov-panrep.htm.

control for the average educational attainment of fathers (of all ages) in the county when the woman was 17 years old. If new colleges were more likely to be built in areas with high demand for college education, then part of this latent demand for college education by women will be captured by measures of the educational attainments of men.

Finally, we try to assess the magnitude of the bias introduced by endogenous mobility of women using data from the 1980 and 1990 Census Public Use Micro Samples (PUMS). We are worried that some women move to areas with higher college availability in order to attend college. The potential for endogenous mobility is probably the main limitation of the instrument. We try to replicate our first stage regression in the Census sample using information on enrollment status of women 19-21 and the location of residence 5 years prior to the Census. The information on residence 5 years prior to the Census allow us to assign women the number of colleges that were available in the location of residence at age 14-16, that is *before* women reached college age. We regress enrollment status on this measure and compare the results with those from a regression of enrollment status on the number of colleges available when the woman was 14-16 in the current location of residence. The location of residence at age 14-16 is less likely to be contaminated by endogenous mobility than current location of residence.

d) Comparison of OLS and IV Estimates

If education is endogenous, and mothers of better "quality" tend to have higher education, OLS will tend to overestimate the true effect of schooling on birth outcomes. Hence, OLS estimates will be larger than valid IV estimates. On the other hand, there are at least two reasons why IV may exceed OLS.

First, the marginal benefit of schooling for individuals whose education has been affected by college openings may be larger than the average benefit for the population. Our IV estimates reflect the effect of education for women who would not have gone to college had it not been for the fact that a college opened in their county of residence. These women are likely to come from disadvantaged backgrounds relatively to those who were not constrained by college location. Card (2000, 2001) reviews a series of recent studies that measure the effects of education on earnings using instrumental variables such as compulsory schooling laws, differences in the accessibility of

schools, and other institutional features of the education market. These studies typically find that instrumental variables estimates *exceed* OLS estimates of the returns to education, suggesting that the marginal returns to education among the groups affected by changes in the instrument are higher than the average returns in the population.

This is relevant when estimating the effect of schooling on child quality. The women affected by college openings may have high marginal returns to schooling, both in terms of earnings and the adoption of healthier behaviors, so that we may find IV estimates that are larger than OLS. Some of the specification tests described in the previous section shed light on whether women's educational choices appeared to be constrained by lack of college availability.

A second potential reason for IV to exceed OLS is the existence of spillovers. It is possible that increases in educational attainment have spillover effects on other women if, for example, pregnant women talk to other pregnant women and new mothers about infant health . Alternatively, spillover effects could work more indirectly, by changing the standard of care practiced by doctors and hospitals (c.f. Aizer and Currie, 2002). That is, if educated women demand that certain practices are adopted, less educated women may also benefit from those practices.

Lochner and Moretti (2001) show that spillovers can provide an alternative explanation for IV estimates that are high relative to OLS. If an individual's decision to engage in healthy behaviors or to buy health inputs depends on average education levels or the average behavior of other individuals in their cohort and county, IV (using county-cohort level instruments as we do) will estimate the combined effect of own education on outcomes as well as the effect of average cohort education on outcomes. That is, IV will estimate the sum of the individual effect and the spillover effect. If cross-county and cohort variation in average education are small relative to overall variation in education, then OLS will only estimate the individual effect of education. In Appendix B we present a formal analysis of the expected relationship between OLS, IV, and grouped OLS estimates when there are spillovers.

Empirically, we try to determine whether differences between OLS and IV can be attributed to education spillovers by aggregating our data to the county-cohort-level: OLS county-cohort-level estimates which exceed OLS estimates obtained using individual-level data are consistent with the hypothesis that IV exceeds OLS because of spillover effects.

A final reason for IV to exceed OLS is that the OLS estimates are contaminated by measurement error. However, if this were the case, then one might expect to see longitudinal estimates smaller than cross sectional estimates, which we do not find.

6. Estimates from Longitudinal Models

Table 2 presents estimates of the longitudinal model (equation 1) obtained using the sample of mothers who are 16 to 24 years old at first birth, and 17 to 30 years old at second birth. Column 1 of Table 2 presents cross sectional estimates of the effects of maternal education on the probability of low birthweight, prematurity, the use of prenatal care in the first trimester, and marital status, using the pooled first and second births from the panel data set. (Estimation using only the first births produced very similar results). The models are similar to equation (1), except that they are estimated using levels rather than changes in outcomes/education.[18] The estimates suggest that education has a strong effect on all four indicators. For example, an increase in education of one year would reduce the probability of low birthweight by almost 10% relative to the means shown in Table 1.

Column (2) of Table 2 presents estimates from models of the form (1) (changes on changes). The estimates are slightly smaller than those in column (1) and suggest that mothers who increase their education between the first and second births reduce the probability of low birth weight and prematurity as well as increasing their use of early prenatal care and marriage probabilities. An increase in education of one year would reduce the probability of low birthweight by about a half of a percentage point. Relative to the baseline estimates in Table 1, this translates into a reduction of almost 10%. The effect on the probability of preterm birth is smaller, at 5.5%. Some of the improvements in birth outcomes may be mediated through increases in the use of prenatal care, and through the increases in the probability of marriage.

The estimates in column (2) of Table 2 are for the full sample of women 16 to 24 at the time of the first birth. The next two columns divide the sample into those who started with a high school degree at the time of the first birth, and mothers who ended with a high school degree at the time of

[18] As mentioned, data on smoking and husband education are not available in this sample for sufficient years to be examined in this sample, which ends in 1993.

the second birth. The motivation for this division is twofold. First, it is possible that the effect of an extra year of schooling is different depending on the initial level of schooling. For example, the means in Table 1 suggest that an extra year of schooling at low levels of education has a larger impact than an extra year at high levels of education. Second, this division is intended to facilitate the comparison of the panel data estimates with the IV estimates to be presented below, since the IV estimates are based on increases in schooling from relatively high initial levels of education (from high school to community college or to four year college).

The estimates in columns 3 and 4 suggest that the effects of education are significant for both groups, although they are larger for less educated women, as one might expect if decreasing returns to education set in after some point. In particular, the estimated effects of education on low birth weight and the probability of marriage are roughly double in the less educated group.

7. Instrumental Variable Estimates

a) Graphical Analysis

We begin our analysis of college openings by graphically showing how maternal education in the sample of first-time mothers 24 and older changed in the ten years before and after the opening of a four-year college. Figure 2 shows the average years of schooling before and after a four year college opening, after controlling for county dummies and dummy variables for mother's cohort of birth.[19] For purposes of illustration, we select openings where there were no other openings in the same county either 10 years before or 10 years after the year of the index college opening. (In the formal analysis below we use all the openings). The figure shows that, on average, maternal education levels are higher in the 10 years following the opening than they were in the 10 years preceding. Moreover, there is no particular trend in maternal education in the county prior to the opening of a new four-year college.

The four panels of Figure 3 show the same story in more detail. The pattern for four-year college attendance is quite similar to the pattern for average years of education shown in Figure 2.

---

[19] In particular, we regressed years of schooling on dummies for time equal to t-10, t-8, t-6, t-4, t-2, t, t+2, t+4, t+6, t+8, t+10, county dummies and cohort dummies. The figure shows the coefficients on dummies for time equal to t-10, t-8, t-6, t-4, t-2, t, t+2, t+4, t+6, t+8, t+10, where t is the time of college opening. Figures 2 to 4 are obtained in a similar way.

This is consistent with the hypothesis that most of the increase in education that arises from the opening of a four year college comes from increases in the fraction of women with a four year college education. Furthermore, most of the increase in the fraction of women with a four year college degree appears to be coming from a reduction in the number of women who have only a high school degree. This lends some credibility to the hypothesis that the college opening induced some women who would have otherwise stopped studying after graduating from high school to go to college. There is little trend in either the number of women with some college (but no four-year degree) or in the number of high school dropouts.

Detecting pre-trends in outcomes is even more important than detecting pre-trends in education. Figure 4 shows the impact of college openings on the incidence of low birthweight and premature birth. Figure 4 shows that while there were no clear trends in the incidence of low birthweight or prematurity in the 10 years prior to the opening of a new college in the county, there is a drop in the incidence of these negative outcomes following the opening.

Figure 5 is like Figure 4, but includes college openings that occur when the mother is 25. Most women should be unaffected by a college opening that occurs past their college age. As expected, we see no clear effect of these later openings on outcomes.

b) First Stage Estimates of the Effects of College Openings on Maternal Education

We now turn to a more formal analysis of the effect of college openings on schooling. Our first stage results are shown in Table 3. Column (1) of panel 1 shows that an increase of one four-year college per 1,000 persons 18 to 22 would result in almost a full year more of maternal education among women 24 and over at the time of first birth. The mean of this variable in our sample is .082 indicating that at the mean, new college openings increased maternal education by approximately .08 years. The effect of a new two-year college is substantially smaller: An increase of one college per 1,000 young adults would increase schooling by about a fifth of a year. Evaluated at the sample mean of .05 two year colleges per 1,000, this coefficient implies that increases in the availability of two-year colleges increased maternal schooling by about .01 years.[20]

[20] A second way to interpret these estimates is to ask how many more college-educated women resulted from the opening of, for example, a new four-year college? Model 2 of Table 3 indicates that such an opening increased the probability that a first-time mother 24 or older had 4 years of college by about 20 percentage points. Since the average county has approximately 10,000 18 to 22 year olds, this implies that one new four year college would increase the

Column 2 shows that adding state-cohort trends to the model has little effect on the estimates. Column 3 indicates that controlling for the average education of fathers in the county when the mother was 17 reduces the coefficients–in this specification a new four year college (per 1,000) increases maternal schooling by .7 rather than .95 years--but the effect remains sizeable and statistically significant.

The remaining panels of Table 3 show how the change in educational attainment break down. A new four-year college (per 1,000) increases the probability of obtaining a four-year degree by 19 percentage points, but has little effect on the probability that only "some college" is completed. A new two-year college (per 1,000) increases the probability of obtaining a four-year degree by 2.5 percentage points, but has a larger effect on the probability of some college, increasing it by 3.2 percentage points. Like the figures, model 4 indicates that most of the increases in schooling following a college opening are coming from a decrease in the number of women obtaining only a high school degree.

Estimates in panels 2 to 5 of Table 3, and in particular the comparison of model 2 with model 3, are inconsistent with the hypothesis that colleges open in areas where education is increasing (or is expected to increase) at all levels. Model 2 and model 3 suggest that the opening of a four or two year college affects the education distribution at the "right" level, and has little spillover onto other levels of schooling.[21]

Having shown that on average new colleges appear to be associated with increases in schooling, we now present separate estimates for public and private colleges. As discussed above, if

---

probability that these mothers had a college education by 2 percentage points. Relative to the baseline of 42%, this is an increase of about 5% in the probability of college. Since the average county in our sample has 165 first time moms 24 to 45, a 5% increase would imply that approximately 8 more of women per county per year received a college education as a result of the college opening.

[21]We have also estimated first stage models separately for three different cohorts of mothers: Those who were 17 between 1940 and 1960, those who were 17 between 1960 and 1980, and those who were 17 between 1980 and 2000 (results are not reported in the table). In Table 3, the coefficient on four-year colleges was .95. The comparable coefficients for the effect of four-year colleges on the three cohorts are: 1.19, .903, and 1.53 (all statistically significant). The comparable coefficients for the effects of two-year colleges are: .007, .173, and .214 (but only the middle coefficient is statistically significant). Thus the effect of having a four-year college open in ones own county is not noticeably weaker for the younger cohorts than for the older ones, whereas the effects of two-year colleges are greatest for those cohorts who experienced the large boom in two-year college construction. We speculate that part of the reason for this pattern may be the increase in the average size of colleges over the period under consideration.

college openings in a woman's county affect maternal education by reducing the costs of attendance, then we might expect to see a larger effect for public colleges than for private colleges, given the usual differences in tuition. Table 4 shows that this is indeed the case. The effect of a new four-year public college is over three times greater than the effect of a new four-year private college. Furthermore, this finding, together with the fact that the location of public colleges is often compensatory, is inconsistent with the hypothesis that private colleges tend to locate in areas with higher expected increases in schooling. There are not enough private two-year colleges to perform the same test on two-year colleges.[22]

Table 5 shows alternative specifications of our first stage equation (2), in which measures of college availability when the mother was 25 or older are added to the base specification shown in Table 3. Finding that college availability when the mother was 25 or older has a *positive* impact on schooling would cast doubt on our estimates. The coefficients on college availability at age 17 remain generally similar to those in Table 3, confirming that this measure of college availability has a positive effect on maternal education. Surprisingly, the coefficients on college at later ages in column 1 are negative. However, column 3 shows that once state-mother cohort time trends and fathers' average educational attainment when the mother was 17 are added to the model, not a single one of the measures of college availability at older ages is statistically significant. We perform an F-test to test whether the coefficients on college availability at 17 are equal to the coefficients on college availability at later ages, and we always reject this restriction. Thus, Table 5 provides evidence that it is college availability at 17 rather than at older ages that matters. This suggests that our results regarding the effects of college openings on maternal education are not driven by underlying trends in college enrollments or by the migration of women older than 24 seeking jobs (or even husbands) in the county of the new college. We present further specification checks in Section c below.

b) OLS and IV Estimates of the Effect of Maternal Education

OLS estimates of the effects of maternal education on infant health, marriage, husband's

[22]We have examined the sensitivity of our IV results to the exclusion of private 4-year colleges (most 2-year colleges are public). We found that we could not reject the hypothesis that the estimates were the same whether or not these colleges were excluded.

education, probability of smoking and parity are shown in column 1 of Table 6 for our sample of older mothers. These estimates are very similar to those in Table 2. For example, they suggest that an additional year of maternal education reduces the probability of low birth weight by 0.5 percentage points (or ten percent), and lowers the risk of prematurity by 0.44 percentage points (or 6 percent).

In this sample, we can examine a wider range of outcomes, including the education of the husband, parity, and the incidence of smoking. Table 6 suggests that much of the positive impact of higher education on birth outcomes may be coming through reductions in smoking: An additional year of education reduces the probability of smoking by more than thirty percent. The increases in four and two-year colleges are estimated to have reduced smoking by three percent and two percent, respectively.

Effects on the use of prenatal care may be interpreted either as the result of higher income, or as a change in behavior. In any case, an additional year of education increases the probability that prenatal care began in the first trimester by one percentage point. However, marriage and prenatal care utilization rates are high for all groups in our sample, which means that there is less scope for education to affect these outcomes than smoking. We find that an additional year of education increases the probability of marriage by one percentage point and increases husband's education by .6 of a year.

Instrumental variables estimates are shown in column 2 of Table 6. We do not present IV estimates for husband's education, since it is not plausible that our instruments affect women's education but not men's. All IV estimates have the same sign as the OLS estimates, and are statistically different from zero. To help in interpreting the magnitude of the estimated effects, consider that the increase in maternal education between the cohort of women who went to college in the 1940s and the 1950s and the cohort of women who went to college in the 1980s is about 1.6 years. The probability of low birth weight and pre-term birth decreased by 6 percentage points and 3 percentage points, respectively. Our estimates suggest that 12% of the decrease in the probability of low birth weight and 20% of the decrease in the probability of pre-term birth can be attributed to increased maternal education.

It is striking that the IV estimates are larger than the OLS estimates for all outcomes except

marriage and parity, where the IV estimates are lower than the OLS. One interpretation is that educated women have unobserved characteristics which also make them more marriageable and likely to have fewer children. Controlling for these characteristics using instrumental variables reduces the estimated effect of education on marriage and parity but does not eliminate it. On the other hand, IV estimates of the effects of education on prenatal care utilization and birth outcomes exceed OLS estimates which suggests either that there are large spillover effects of education on these outcomes or that there is a higher than average return to educating the marginal woman in terms of infant health.

To try to shed some light on which of these two explanations is more important, the third column of Table 6 shows estimates of equation (2) using data aggregated to the county*child's year of birth level. As discussed above, if spillovers are important, then we expect these estimates to significantly exceed those in column (1). And if spillovers are to explain the fact that IV estimates exceed OLS estimates, then the estimates in column (3) should resemble those in column (2). On the basis of these comparisons, the outcome for which spillovers appear to be most important is the timely use of prenatal care.

Table 7 presents some evidence regarding the robustness of our OLS and IV results to various changes in specification. The odd columns present estimates which control for state-mother cohort effects, while the even columns also control for average fathers' education when the mother was 17.[23] Table 7 shows that our results are remarkably robust to these changes in specification.

As another check intended to probe the robustness of our findings, in Table 8 we show estimates for a sample that includes mothers 22 to 45. The sample size increases by 30%. As argued in Section 4, this sample has the advantage of being more representative, because many mothers have their first baby before 24. On the other hand, it has the disadvantage that many women are still in college at 22. The first stage and second stage results are generally similar to the ones reported in Table 3 and 5. It is interesting to note that the IV estimates of the effect of schooling on infant health are closer to the OLS estimates than the corresponding IV estimates in Table 6 (rows 1 and 2).

---

[23]We also estimated alternative models that included region*cohort dummies in addition to the state*cohort trends. These specifications allow cohort effects to vary over regions in a non-linear way, while the state-cohort time trends constrain differences between state to be linear. In any case, the estimates including region*cohort effects were very similar to those shown in Table 7.

d) Additional IV Specification Checks

We now turn to the additional specification checks described in Section 5c. We begin by considering single-sex colleges. If our identification strategy is valid, then the transformation of a male only college into a coed college should increase maternal education, while the transformation of a female only college into a coed college should have no effect (or a negative effect if women are crowded out) on maternal education. In our sample, there are 104 cases of four-year colleges switching from male-only to coed, and 85 cases of four year colleges switching from female-only to coed. Virtually all two-year colleges began as coed institutions. Table 9 shows that having a college change from being male-only to being coed, has an effect on maternal education that is remarkably similar to the effect of a new four year college opening shown in Table 3. The effect of having a college change from being female only to coed is smaller, and not statistically significant.

In Table 10 we ask whether the effects of new college openings are greater in locations far from other colleges. In the first specification (Model 1), we divide counties into two groups according to whether the number of colleges (normalized by cohort size) in adjacent counties is above or below the median. These variables are then interacted with our measures of two and four-year college availability. The results indicate that new four-year colleges do have a greater effect in college-poor areas, but the difference between college-rich and college-poor areas is not statistically significant. On the other hand, new two-year colleges are found to have a much greater impact on maternal education in college-poor areas.

In a second specification (Model 2 of Table 10), we adopt a continuous measure of the availability of colleges in adjacent counties, defined as the average number of colleges in adjacent counties (again normalized by the relevant cohort sizes). This specification suggests that the effect of a new college falls off rapidly with the number of colleges in contiguous areas: An increase of one four-year college per 1,000 residents 18 to 22 is estimated to increase maternal education by almost a year in areas with no colleges in contiguous counties. But the effect falls to zero as the number of colleges in contiguous counties rises to one per 1,000 residents in the relevant age cohort.

e) Endogenous Mobility

Finally, we turn to the issue of endogenous mobility. There are two types of endogenous mobility that are potentially problematic for our instrumental variables strategy. First, 17 or 18 year old women may move to counties that experience college openings to attend college. Second, college educated women may move to counties that experience college openings after they graduate from college but before they have their first child. In Table 5 we presented evidence that indicates that the second type of mobility does not appear to be very significant. Mobility of the first type is more worrisome. It is undoubtably the case that some women move to areas with higher college availability in order to attend college. The question is whether such mobility is important enough to account for a significant portion of the documented correlation between number of colleges and maternal education (i.e. our first stage coefficient). If mobility explains a significant portion of our first stage coefficient, then our instrument is invalid. We want to make clear that, because of data limitations, we cannot completely rule out this possibility. We do have some indirect evidence that we hope may shed some light on the magnitude of the problem.

First, consider that new colleges in our sample appear to be, for the most part, non-selective institutions that are unlikely to attract a significant portion of the student body from far away. This is, of course, just a conjecture, since data on the county of residence of the students prior to their enrollment do not exist. The only available data report the state of residence prior to enrollment. In Table 11 we report the percentage of college students who are from the state where the college is located, by the age of the college, in 1998. The Table confirms that, while old, more established institutions attract a significant number of out of state students, new, less established colleges attract mostly in state students. For example, while more than a third of the students enrolled in four year colleges that are 100 years old or older are from out of state, only 9 percent of students enrolled in colleges 9 years or younger are from out of state. The corresponding figures for two year colleges are even lower. Unsurprisingly, community college students are overwhelmingly local, irrespective of the age of the college.

Although suggestive, these numbers are less than ideal. On one hand, the figures shown are for 1998. Since student mobility has steadily increased since 1940 (Hoxby, 1996), mobility in most of the years we consider is likely to be even lower. On the other hand, Table 11 reports student

mobility across states, and is likely to underestimate mobility across counties.

For a second piece of indirect evidence, we use the 1980 and 1990 five percent Public Use Micro Samples of the Census and try to assess the effect of the instrumental variable on college enrollment using the county of residence at ages younger than college age. We compare the estimated effect with the effect of the instruments on college enrollment using current county of residence. Specifically, we take all white women 19 to 21 who report information on the location of residence five years prior to the Census (that is, when the respondent was 14 to16). We then regress a dummy equal to one if the respondent is enrolled in school on the (normalized) number of four and two year colleges that existed when the respondent was 14 to16 in the location where she lived when she was 14 to16. We also regress the enrollment dummy on the (normalized) number of four and two year colleges that existed when the respondent was 14 to16 in the location where she currently lives. Both models include year and location fixed effects.

Ideally, we would like to replicate our first stage regressions using Census data to check whether estimates differ when county of residence at age 14-16 is used instead of current county of residence. However, data limitations preclude an exact replication. First, the dependent variable is enrollment, not schooling achievement, because the individuals for whom we have information on county of residence at age 14-16 may still be in college at the time of the Census. In other words, while the Natality IV sample includes women 24 or older, the Census sample includes women 19 to 21 years old. Second, we include all women, not only mothers, since we have no way of knowing which of the women will end up having children. Third, the smallest geographical identifier in the Census is the PUMA, not the county. PUMA are smaller than counties in urban areas, and larger than counties in rural areas. For example, for the nine-county San Francisco Bay Area, there are 48 PUMAs. We aggregate PUMAs in urban areas to the level of metropolitan area. See the Data Appendix text for more details. Finally, the Census reports the information on PUMA of residence five years ago only in 1980 and 1990. For these reasons, Census estimates are not exactly comparable with the first stage estimates in Table 3.

Estimates in the top panel of Table 12 indicate that the coefficient on four year colleges that existed when the respondent was 14-16 in the location where she lived at age14-16 is 0.104, indicating that an increase in the (normalized) number of four year colleges in the location of

residence *before* the respondent reaches college age is associated with an increase in the probability of enrollment. Estimates in the bottom panel indicate that the coefficient on four year colleges that existed when the respondent was 14-16 in the location where the respondents is at the time of the Census is 0.106, not very different from the corresponding coefficient in the upper panel. The location of residence *before* the respondent reaches college age is less likely to be contaminated by endogenous mobility than current location of residence. The similarity of the two estimates lends some credibility to the assumption that endogenous mobility is not solely responsible for our first stage results for four years colleges, although the difference between Natality and the Census samples precludes firm conclusions.

The similarity between upper and lower panel estimates does not extend to two year colleges. The coefficient for two year colleges in the upper panel is 0.011. The fact that the coefficient for two year colleges is smaller than the corresponding coefficient for four year colleges is consistent with results in Table 3. However, the coefficient for two year colleges in the lower panel is negative. This last finding is surprising and stands in sharp contrast both with the coefficient in the upper panel and the first stage coefficient in Table 3. We don't have a definitive explanation for it. But it should be noted that only a very limited number of new two years colleges enter the sample between 1980 and 1990. Most of the new two years colleges in the sample open during the 1960s and 1970s (Figure 1).

In summary, our estimates of the effects of college openings on maternal education suggest that four year colleges have a greater effect on four year college completion than two year colleges (and vice versa); that the education of women who are 25 or older when a college opening occurs is not affected by the opening; that female-only colleges increase women's education, but male-only colleges do not; that public colleges have a greater impact than private colleges; and that college openings in "college poor" locations have larger effects. All these findings suggest that the estimated effects of college openings are unlikely to simply reflect an omitted third factor. Moreover, our Census results suggest that increases in the number of college educated women following the opening of a four year college are unlikely to be *solely* due to inflows of women seeking education. Thus, while it is impossible to definitively prove the validity of our instrumental variables strategy, the specification checks provide some additional evidence supporting our identification assumptions.

7. Conclusions

We provide new evidence regarding the effect of maternal education on infant health, and on a series of other factors that are likely to influence infant health including smoking, use of prenatal care, marriage, and fertility behavior. Our estimates suggest that increases in maternal education over the past 30 years have had large positive effects on birth outcomes. We estimate that an additional year of education reduces the incidence of low birth weight by approximately ten percent, and reduces the incidence of pre-term birth by six percent, on average. These effects arise because education affects maternal behavior (by reducing smoking by more than 30 percent, for example); it increases earnings, possibly relaxing budget constraints; it improves women's marriage markets; and it reduces fertility. As expected, the estimated effects are larger for less educated women, but remain substantial even among high school graduates and college graduates. We also find some indirect evidence of educational spillovers, particularly in the use of prenatal care in the first trimester.

Although we cannot rule out alternative explanations—in particular, we cannot completely rule out endogenous mobility—the qualitative results are robust to the use of longitudinal data and instrumental variables methods. The fact that longitudinal models and instrumental variable estimates depend on very different assumptions lends some credibility to our findings.

In addition, our first stage estimates are of interest in their own right, since they demonstrate that the boom in the construction of new colleges during the 1960s and 1970s had a significant impact on the education of mothers. Specifically, the mean county experienced an increase of .08 four-year colleges and .05 two-year colleges per 1,000 young people 18 to 22. These improvements in the education infrastructure increased average maternal schooling by .08 years and .01 years for four and two-year colleges respectively.

On average, a change in .09 years of education reduced low birthweight by about one percent, and reduced the incidence of preterm birth by about half of one percent, according to our OLS and panel data estimates. If we use the IV estimates to calculate the impact on those most likely to have been affected, then the increase in education induced by the college openings is estimated to have reduced the incidence of low birthweight and preterm delivery by closer to two percent and one

percent, respectively. While these may seem like small improvements, the costs of low birthweight and prematurity are large. For example, it is estimated that between birth and age 15, low birthweight children incur an additional $5.5 to $6 billion more in health, education, and other costs than children of normal birthweight (March of Dimes, 2002).

Traditionally, research on the benefits of schooling has focused on the *private* returns to education, i.e. the effect of a worker's education on her own wage. There is growing evidence that the *social* return to education may exceed the private return.[24] Improved birth outcomes are another benefit of education that is not reflected in the wages of educated individuals. Hence, our results suggest that estimates of the returns to education which focus only on increases in wages significantly understate the social return. Moreover, if educating mothers improves the health, educational attainment, and labor market outcomes of children, then this is a significant inter-generational benefit.

---

24 Education has been shown to benefit productivity, innovation, and lower crime rates. See for example Moretti 2000; and Lochner and Moretti 2002.

References

Becker, Gary S. A Treatise on the Family (Cambridge: Harvard University Press) 1981.

Becker, Gary S. "Human Capital, Effort, and the Sexual Division of Labor" Journal of Labor Economics 3, January 1985, s33-s58.

Becker, Gary S. and H. Gregg Lewis. "On the Interaction Between the Quantity and Quality of Children", Journal of Political Economy, 81:2, 1973, S279-288.

Behrman, Jere R, and Mark R. Rosenzweig. "Does increasing women's schooling raise the schooling of the next generation?", American Economic Review, 92:1, March 2002, 323-334.

Blau, Francine and Adam Grossberg. "Maternal Labor Supply and Children's Cognitive Development", Review of Economics and Statistics, 74(3), August, 1992.

Card, David. "The Causal Effect of Education on Earnings" in The Handbook of Labor Economics, v3, Orley Ashenfelter and David Card (eds.), (New York: North Holland) 2000.

Card, David. "Estimating the return to schooling: Progress on some persistent econometric problems", Econometrica, 69:5, Sept. 2001, 1127-1160.

Card, David and Thomas Lemieux. "Can falling supply explain the rising return to college for younger men? A cohort-based analysis", Quarterly Journal of Economics, 116:2, May, 2001, 705-746.

Currie, Janet. "Child Health in Developed Countries" in The Handbook of Health Economics Anthony Culyer and Joseph Newhouse (eds). (New York: North Holland) 2000.

Currie, Janet and Rosemary Hyson. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status?: The Case of Low BirthWeight", American Economic Review, May 1999.

Desai, Sonalde, Lindsay Chase-Lansdale and Robert Michael. "Mother or Market? Effects of Maternal Employment on the Intellectual Ability of 4-Year Old Children", Demography, 26, 1989.

Desai, Sonalde and Soumya Alva. "Maternal education and child health: Is there a strong causal relationship?" Demography 35:1, Feb. 1998, 71-81.

Fuchs, Victor. Economic Aspects of Health (Chicago: University of Chicago Press for NBER) 1982.

Goldin, Claudia. "The Meaning of College in the Lives of American Women: The Past One-Hundred Years", NBER Working Paper #4099 (Cambridge MA: National Bureau of Economic Research), June 1992.

Goldin, Claudia. "Career and Family: College Women Look to the Past", in Francine Blau and Ronald Ehrenberg Gender and Family Issues in the Workplace (New York: Russell Sage Press) 1997.

Goldin, Claudia and Larry Katz, "The Shaping of Higher Education: The Formative Years in the United States, 1840 and 1940", Journal of Economic Perspectives, vol 13 37-62 (1999)

Grossman, Michael. The Demand for Health: A Theoretical and Empirical Analysis, PhD Dissertation, Columbia University, 1970.

Grossman, Michael. "The Human Capital Model of the Demand for Health" in Handbook of Health Economics Anthony Culyer and Joseph Newhouse (eds.) (New York: North Holland) 2000.

Harworth, Irene, Mindi Maline, Elizabeth DeBra. "Women's Colleges in the United States: History, Issues and Challenges", http://www.ed.gov/offices/OERI/PLLI/webreprt.html, August 19, 2002.

Hoxby, Caroline. "The Effects of Geographic Integration and Increasing Competition in the Market for College Education", NBER Working paper 6323, Dec. 1997.

Kane, Thomas and Cecilia Rouse. "Labor-Market Returns to Two and Four-Year College", American Economics Review 85:3, Jane 1995, 600-614.

Lam, David. "Marriage Markets and Assortative Mating with Household Public Goods: Theoretical Results and Empirical Implications", Journal of Human Resources 23:4, Fall 1988, 462-487.

Lazear, Edward. "Intergenerational Externalities", Canadian Journal of Economics, v16 #2 (May 1983, 212-228.

Lightwood, James, Cairan Phibbs, and Stanton Glantz. "Short-Term Health and Economic Benefits of Smoking Cessation: Low Birth Weight", Pediatrics, 104:6, December 1999, 1312-1320.

Lleras-Muney, Adriana. "The Relationship Between Education and Adult Mortality in the U.S.", Princeton University Dept. of Economics, xerox, May 2001.

Lochner, Lance and Enrico Moretti. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports", *NBER Working Paper 8606*, 2001.

March of Dimes. Health Library: Low Birthweight/Prematurity www.modimes.org/HealthLibrary /355-1477.htm, March 2002.

Mare, Robert. "Five Decades of Educational Assortive Mating", American Sociology Review 56, 1991, 15-32.

Marshall, Alfred. Principles of Economics (New York: Macmillan) 1890.

Moretti, Enrico. "Estimating the Social Return to Education: Evidence from Longitudinal and Repeated Cross-Sectional Data", forthcoming in Journal of Econometrics.

Neidell, Matthew. "Early Time Investments in Children's Human Capital Development: Effects of Time in the First Year on Cognitive and Non-Cognitive Outcomes", Dept. of Economics UCLA, xerox, October 2000.

Oppenheimer, Valerie. "A Theory of Marriage Timing: Assortive Mating Under Varying Degrees of Uncertainty", American Journal of Sociology, 94, 563-91.

Parcel Toby and Elizabeth Menaghan. Parent's Jobs and Children's Lives (New York: Aldine de Gruyter) 1994.

Rindfuss, Ronald, S. Philip Morgan and K. Offutt. "Education and the Changing Age Pattern of American Fertility: 1963-1989", Demography, 33:3, August 1996, 277-290.

Rosenzweig, Mark and T. Paul Schultz. "The Behavior of Mothers as Inputs to Child Health: The Determinants of Birth Weight, Gestation, and Rate of Fetal Growth," in Economic Aspects of Health, Victor Fuchs (ed.), University of Chicago Press: Chicago, 1982.

Rosenzweig, Mark and T. Paul Schultz. "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight," Journal of Political Economy, 91, October 1983, 723-746.

Rosenzweig, Mark and T. Paul Schultz. "The Stability of Household Production Technology, A Replication," The Journal of Human Resources, 23, Fall 1988, 535-549.

Rosenzweig, Mark R, and Kenneth I. Wolpin. "Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement", Journal of Human Resources, 29:2, Spring 1994, 670-693.

Rouse, Cecilia. "Democratization or Diversion? The Effect of Community College on Educational Attainment", Journal of Business and Economic Statistics, 13:2, April 1995, 217-224.

Ruhm, Christopher. "Parental Employment and Child Cognitive Development" (NBER: Cambridge MA) Working Paper # 7666, April 2000.

Thomas, Duncan, John Strauss, and Maria-Helena Henriques. "How Does Mother's Education Affect Child Height?", Journal of Human Resources, 26:2, Spring 1991, 183-211.

Welch, Finis. "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust", Journal of Political Economy, 87:5, October 1979.

The World Bank. "World Development Report 1993: Investing in Health" (New York: Oxford University Press) 1993.

Figure 1: Number of Four and Two Year Colleges

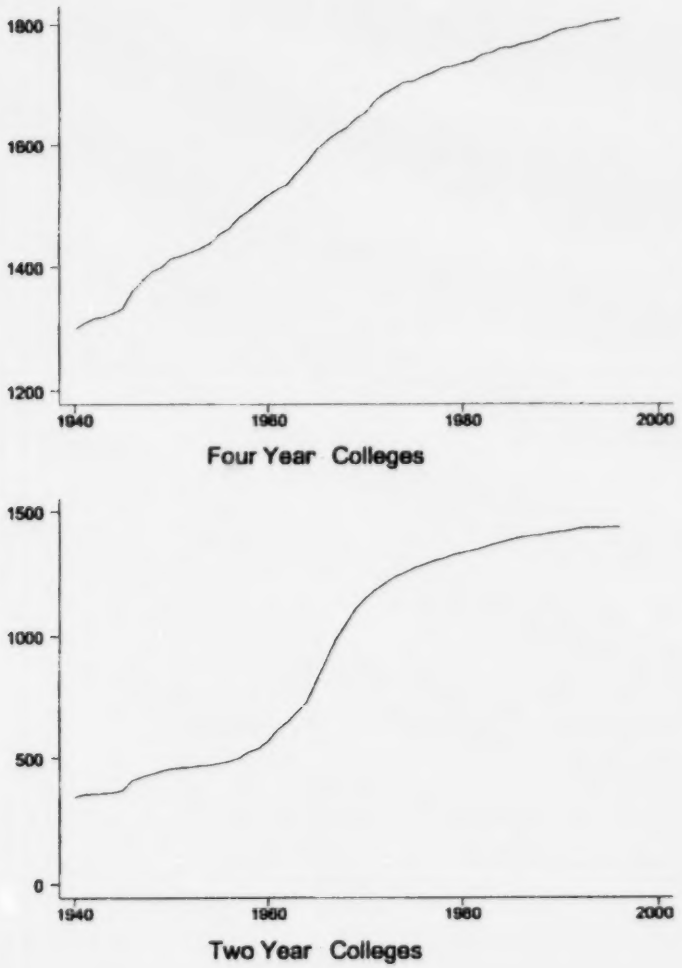
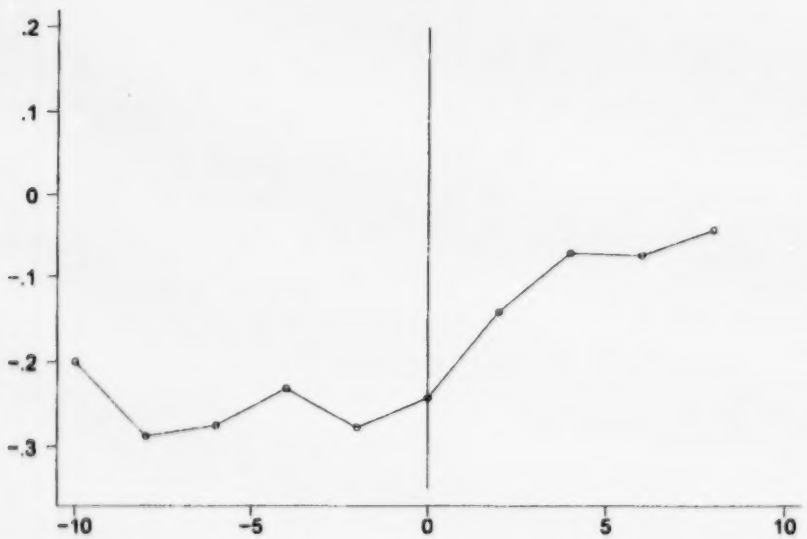
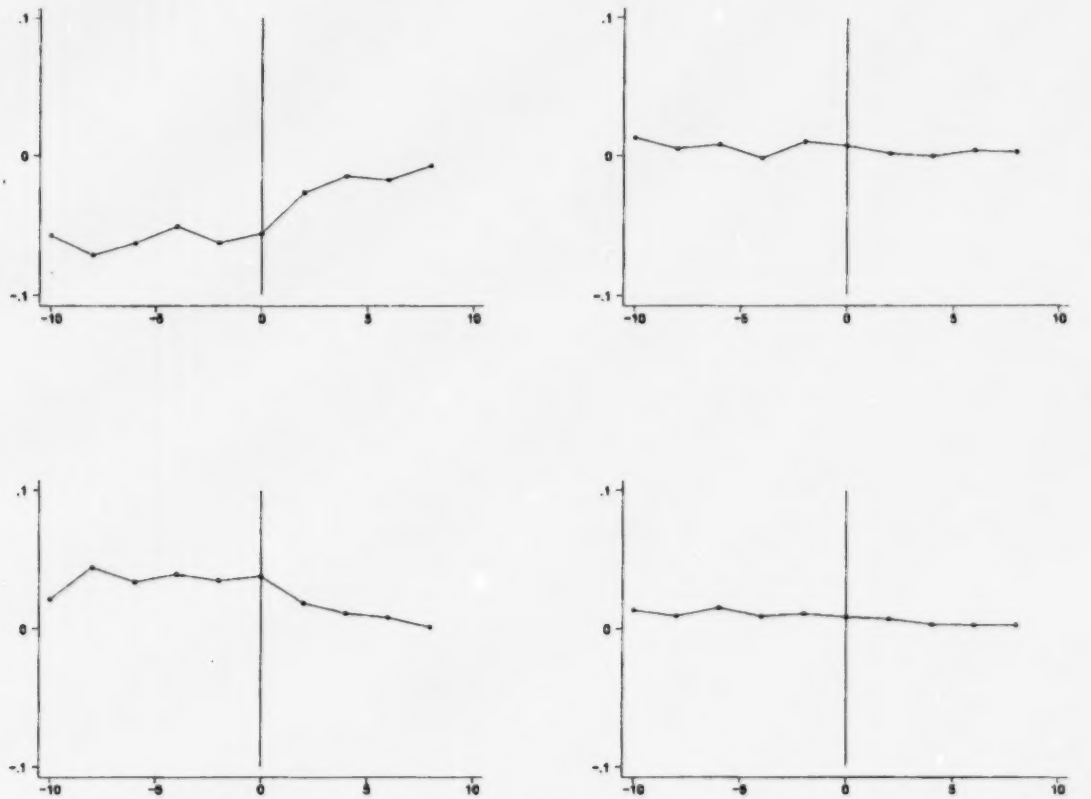


Figure 2: Average Years of Schooling of Women, Before and After the Opening of a Four Year College



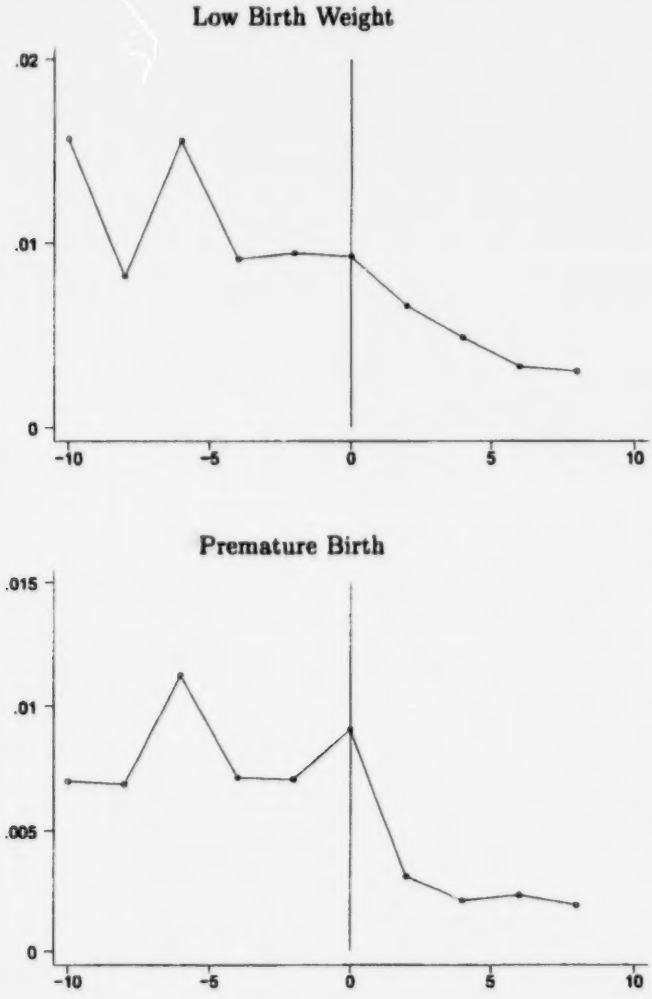
Notes: The Figure shows average years of schooling in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. The figures includes only college openings that occurred at least 10 years apart from other openings in the same county. Time 0 is when the college opens and the mother is 17 years old.

Figure 3: Probability of College, Some College, High School and Less than High School, Before and After the Opening of a Four Year College



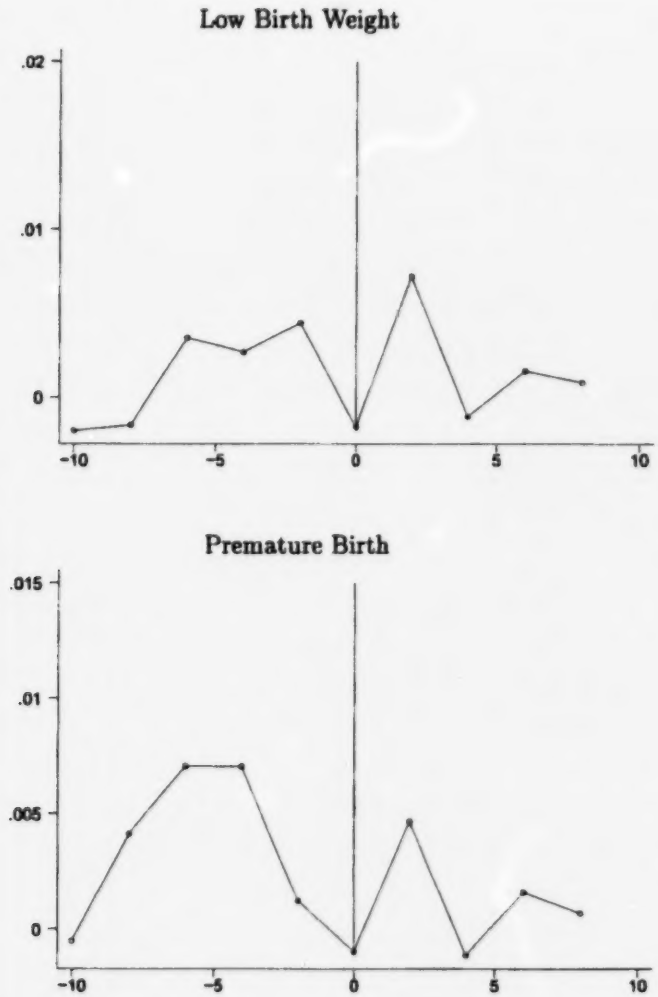
Notes: The top-left panel shows the conditional probability of college education in the years preceding and the years following the opening of a four year college, controlling for county dummies and mother cohort dummies. The top-right panel is for commuity college; the bottom left panel is for high school; the bottom-right panel is for less than high school. Time 0 is when the college opens and the mother is 17 years old. The figures includes only college openings that occurred at least 10 years apart from other openings in the same county.

Figure 4: Average Birth Outcomes, Before and After the Opening of a Four Year College - Mother is 17 When the College Opens



Notes: The top panel shows the conditional probability of low birth weight in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. The bottom panel shows the conditional probability of premature birth in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. In both figures, time 0 is when the college opens and the mother is 17 years old. The figures include only college openings that occurred at least 10 years apart from other openings in the same county.

Figure 5: Average Birth Outcomes, Before and After the Opening of a Four Year College - Mother is 25 When the College Opens



Notes: The top panel shows the conditional probability of low birth weight in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. The bottom panel shows the conditional probability of premature birth in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. In both figures, time 0 is when the college opens and the mother is 25 years old. The figures include only college openings that occurred at least 10 years apart from other openings in the same county.

Table 1: Summary Statistics

	All	Panel Sample		IV Sample				
				All	College	Some College	High School Only	High School Drop Out
	1st births 16-45 (1)	1st births 16-24 (2)	2nd births 17-30 (3)	1st births 24-45 (4)	1st births 24-45 (5)	1st births 24-45 (6)	1st births 24-45 (7)	1st births 24-45 (8)
Mother's Education	12.95 (2.22)	11.87 (1.62)	12.24 (1.60)	14.21 (2.08)	16.3 (0.47)	14.27 (0.44)	12 (0)	9.47 (2.15)
Low Birth Weight	0.054	0.050	0.035	0.049	0.041	0.047	0.057	0.089
Pre-term Birth	0.079	0.075	0.062	0.069	0.062	0.068	0.075	0.100
Prenatal Care	0.827	0.755	0.795	0.921	0.945	0.925	0.894	0.759
Smoked During Pregnancy	.145	--	--	0.078	0.023	0.080	0.168	0.340
Married	.772	0.842	0.923	0.923	0.968	0.930	0.890	0.721
Husband's Education	13.1 (2.31)	--	--	14.20 (2.65)	15.55 (1.65)	14.39 (1.96)	12.88 (1.94)	10.87 (2.7)
Mother Age	23.83 (5.14)	20.29 (2.38)	23.3 (2.81)	28.12 (3.56)	29.02 (3.6)	27.80 (3.1)	27.31 (3.3)	27.30 (3.6)
Year Mother was 17	1980 (8.7)	1976.4 (4.96)	1976.4 (4.96)	1974.8 (7.5)	1975.0 (7.7)	1975.6 (7.4)	1974.1 (7.3)	1973.4 (8.6)
Per Capita Number of Four Year Colleges in County, when Mother was 17	0.0796 (0.0984)	0.0838 (0.129)	0.0838 (0.129)	0.0819 (0.0925)	0.0856 (0.0856)	0.0799 (0.0903)	0.0785 (0.0932)	0.0766 (0.0974)
Per Capita Number of Two Year Colleges in County, when Mother was 17	0.0548 (0.0850)	0.0633 (0.120)	0.0633 (0.120)	0.0502 (0.0696)	0.0492 (0.0492)	0.0519 (0.0672)	0.0497 (0.0739)	0.0514 (0.0850)
Median Family Income in County when Mother was 17	32510 (7150)	27156 (7768)	27156 (7768)	33273 (7467)	34082 (7716)	33670 (7120)	32359 (7176)	29942 (7751)
Percentage County that was Urban when Mother was 17	0.71 (0.27)	0.51 (0.31)	0.51 (0.31)	0.77 (0.24)	0.80 (0.23)	0.78 (0.23)	0.74 (0.26)	0.70 (0.29)
Sample Size	20,422,042	478,272	478,272	671,468	279,574	169,916	204,394	17,584
Parity (here we include all births rather than 1st births)	2.0 (1.3)	1 (0)	2 (0)	2.4 (1.2)	2.0 (1.1)	2.3 (1.2)	2.5 (1.2)	3.4 (1.2)

Notes: Column 1 reports summary statistics for all first time white mothers 16 to 45. Columns 2 and 3 report summary statistics for white mothers who are 16 to 24 at their first birth and 17 to 30 at their second birth and that are matched across births. Columns 4 to 8 report summary statistics for first time white mothers 24 to 45. For columns 4 to 8, a 10% sample was randomly selected. Averages in the last row include all births, not just first births. Not all variables are available in all years. For smoking, N=166,183. For husband's education, N-= 486,255. Standard deviations in parentheses.

Table 2

Longitudinal Models: The Effect of Changes in Maternal Education on Changes in Infant Health, Prenatal care and Marital Status

	Cross Section	Changes on Changes		
	All Mothers	All Mothers	Mothers with High School Degree Or More at First Birth (Educ. Range: 12-16+)	Mothers with High School Degree Or Less at Second Birth (Educ. Range: 0-12)
	(1)	(2)	(3)	(4)
1. Low Birth Weight	-0.0057	-0.0049	-0.0034	-0.0057
	(0.0003)	(0.0005)	(0.0007)	(0.0008)
2. Pre-Term Birth	-0.0033	-0.0032	-0.0031	-0.0035
	(0.0004)	(0.0006)	(0.0009)	(0.0010)
3. Prenatal care	0.0188	0.0170	0.0152	0.0179
	(0.0006)	(0.0010)	(0.0013)	(0.0016)
4. Marital Status	0.0216	0.0293	0.0162	0.0406
	(0.0005)	(0.0007)	(0.0008)	(0.0011)

Notes: Standard errors in parentheses. The coefficients reported in column 1 are the coefficients on mother's education in a cross-sectional regression. The coefficients reported in columns 2 to 4 are the coefficients on changes in mother education. The sample in columns 1 and 2 is the same. The sample was obtained as follows: first, women with parity =1 were selected. Then, women with parity = 2 were matched to women with parity =1 based on county, date of birth of first child (month and year), mother's age at first birth, mother's year of birth and mother's state of birth. Finally, only cells with a unique observation in both years were kept (one-to-one matches). The sample is therefore an actual panel of mothers with two children, where the same mother is observed at the time of the first birth and the time of the second birth. The sample includes mothers who are 16-24 at their first birth and 17-30 at their second birth. The sample used in column 3 includes mothers who have high school or more at first birth. The sample used in column 4 includes mothers who have high school or less at second birth. Smoking is missing because it was reported first in 1989, while the month of birth for the last birth was reported only until 1993. All models include unrestricted mother age at first birth effects, unrestricted mother age at second birth effects and unrestricted mother decade of birth effects. See text for details.

Table 3
The Effect of College Openings on Maternal Education

		(1)	(2)	(3)
Model 1: Years of Schooling				
	Four Year Colleges	0.950	0.952	0.716
		(0.046)	(0.046)	(0.042)
	Two Year Colleges	0.176	0.158	0.145
		(0.046)	(0.046)	(0.046)
	R squared	0.11	0.12	0.12
Model 2: 4 Years or More of College				
	Four Year Colleges	0.198	0.199	0.150
		(0.010)	(0.010)	(0.009)
	Two Year Colleges	0.025	0.021	0.018
		(0.010)	(0.010)	(0.010)
	R squared	0.10	0.10	0.10
Model 3: Some College				
	Four Year Colleges	0.000	0.046	0.000
		(0.006)	(0.006)	(0.006)
	Two Year Colleges	0.032	0.032	0.032
		(0.007)	(0.007)	(0.007)
	R squared	0.02	0.03	0.03
Model 4: High School Only				
	Four Year Colleges	-0.156	-0.156	-0.120
		(0.009)	(0.009)	(0.009)
	Two Year Colleges	-0.053	-0.051	-0.051
		(0.010)	(0.010)	(0.010)
	R squared	0.07	0.07	0.07
Model 5: Less Than High School				
	Four Year Colleges	-0.026	-0.026	-0.026
		(0.003)	(0.003)	(0.003)
	Two Year Colleges	-0.003	-0.003	-0.003
		(0.004)	(0.004)	(0.004)
	R squared	0.03	0.03	0.03
State-Mother Cohort Trends			Yes	Yes
Men's Average Education				Yes

Notes: Standard errors in parentheses. All models include mother age effects, mother cohort of birth effects, county*year effects (where year is year of birth of the child), median county income and percent urban in year when mother was 17. "Four year" is the number of four year colleges existing in a county in the year when the mother was 17 normalized by the size of the mother's cohort (number of 17-20 year olds in the county, in thousands). "Two year" is the number of two year colleges existing in a county in the year when the mother was 17 normalized by the size of the mother's cohort. The sample includes all white mothers 24 to 45 with parity equal to 1. See text for details.

Table 4
The Effect of College Openings on Maternal Education: Separate Estimates for Public and Private Four-year Colleges

	Coefficient on Four Year Colleges (1)	Coefficient on Two Year Colleges (2)
Public Colleges	1.814	
	(0.075)	0.181
		(0.046)
Private Colleges	0.503	
	(0.053)	

Notes: Standard errors in parentheses. All coefficients come from one regression. The model includes unrestricted mother age effects, unrestricted mother cohort of birth effects, county*year effects (where year is year of birth of the child), median county income in year when mother was 17, and percent of the county that is urban in year when the mother was 17. Number of colleges is normalized by cohort size. The sample includes all white mothers 24 to 45 with parity equal to 1. There are not enough private two-year colleges to perform the same test on two-year colleges. See text for details.

Table 5

The Effect of College Openings at Ages Later than 17 on Maternal Years of Schooling

	(1)	(2)	(3)
Model 1: Add colleges at Age 25			
Four Year Colleges at Age 17	1.259	1.330	0.941
	(0.161)	(0.161)	(0.153)
Two Year Colleges at Age 17	0.317	0.299	0.191
	(0.104)	(0.103)	(0.102)
Four Year College at Age 25	-0.371	-0.454	-0.264
	(0.180)	(0.178)	(0.171)
Two Year College at Age 25	-0.180	-0.177	-0.057
	(0.105)	(0.105)	(0.107)
Effect at age 17 = Effect at age 25 (p-value)	0.000	0.000	0.000
Model 2: Add colleges at Age 30			
Four Year Colleges at Age 17	1.304	1.304	0.857
	(0.139)	(0.140)	(0.131)
Two Year Colleges at Age 17	0.279	0.241	0.071
	(0.100)	(0.100)	(0.098)
Four Year Colleges at Age 30	-0.444	-0.449	-0.246
	(0.157)	(0.158)	(0.149)
Two Year Colleges at Age 30	-0.122	-0.095	0.063
	(0.100)	(0.100)	(0.100)
Effect at age 17 = Effect at age 30 (p-value)	0.000	0.000	0.000
Model 3: Add colleges at Age 35			
Four Year Colleges at Age 17	1.221	1.234	0.800
	(0.151)	(0.152)	(0.145)
Two Year Colleges at Age 17	0.250	0.236	0.098
	(0.106)	(0.106)	(0.105)
Four Year Colleges at Age 35	-0.325	-0.344	-0.169
	(0.179)	(0.180)	(0.174)
Two Year Colleges at Age 35	-0.184	-0.174	-0.056
	(0.105)	(0.105)	(0.106)
Effect at age 17 = Effect at age 35 (p-value)	0.000	0.00	0.005
Model 4: Add colleges at Age 40			
Four Year Colleges at Age 17	1.120	1.127	0.722
	(0.176)	(0.178)	(0.176)
Two Year Colleges at Age 17	0.197	0.167	0.080
	(0.153)	(0.154)	(0.150)
Four Year Colleges at Age 40	-0.050	-0.064	0.043
	(0.229)	(0.233)	(0.229)
Two Year Colleges at Age 40	-0.247	-0.220	-0.107
	(0.146)	(0.147)	(0.148)
Effect at age 17 = Effect at age 40 (p-value)	0.002	0.002	0.17
State-Mother Cohort Trends		Yes	Yes
Men's Average Education			Yes

Notes: Standard errors in parentheses. All models include mother age effects, mother cohort of birth effects, county*child year-of-birth effects, median county income and percent urban in year when mother was 17. "Four year colleges at age 17" is the number of four year colleges in a county in the year when the mother was 17 normalized by the size of mother's cohort (number of individuals 17-20 year old in the county, in thousands). "Two year colleges at age 17" is the number of two year colleges in a county in the year when the mother was 17 normalized by the size of the mother's cohort. Similar definitions apply for ages 25, 30, 35 and 40. The sample includes all white mothers 24 to 45 with parity equal to 1. See text for details.

Table 6
The Effect of Maternal Education on Infant Health, the Marriage Market and Fertility

	OLS (1)	IV (2)	OLS Grouped (3)
1. Low Birth Weight	-0.0050	-0.0098	-0.0057
	(0.0001)	(0.0038)	(0.0008)
2. Pre-Term Birth	-0.0044	-0.010	-0.0078
	(0.0001)	(0.0044)	(0.0009)
3. Prenatal care	0.0114	0.0234	0.0230
	(0.0001)	(0.0055)	(0.0010)
4. Smoked During Pregnancy	-0.0305	-0.0583	-0.0345
	(0.0004)	(0.0118)	(0.0030)
5. Married	0.0206	0.0128	0.0135
	(0.0002)	(0.0040)	(0.0008)
6. Husband's Education	0.607	--	0.690
	(0.0019)		(0.0081)
7. Parity	-0.121	-0.092	-0.0577
	(0.000)	(0.010)	(0.0030)

Notes: Standard errors in parentheses. All models include unrestricted mother age effects, unrestricted mother cohort of birth effects, county*year effects (where year is year of birth of the child), median county income in year when mother was 17, and percent of the county that is urban in the year when the mother was 17. The instrumental variables in column 2 are the number of four year colleges in a county in the year when the mother was 17 normalized by the size of the mother's cohort (number of 17-20 year olds in the county, in thousands) and the number of two year colleges in a county in the year when the mother was 17 normalized by the size of the mother's cohort. The model in column 3 is estimated on data aggregated to the county-mother cohort of birth level. Models in rows 1 to 6 include all white mothers 24 to 45 with parity equal to 1. Models in row 7 include all white mothers 24 to 45 of any parity. See text for details.

Table 7
The Effect of Maternal Education on Infant Health, the Marriage Market and Fertility: Alternative Specifications

	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
1. Low Birth Weight	-0.0050	-0.0050	-0.0099	-0.0110
	(0.0001)	(0.0001)	(0.0038)	(0.0054)
2. Pre-Term Birth	-0.0044	-0.0046	-0.010	-0.010
	(0.0001)	(0.0001)	(0.0044)	(0.0062)
3. Prenatal care	0.0116	0.0117	0.0241	0.0251
	(0.0001)	(0.0001)	(0.0054)	(0.0077)
4. Smoked During Pregnancy	-0.0305	-0.0365	-0.0623	-0.0607
	(0.0004)	(0.0004)	(0.0118)	(0.0204)
5. Married	0.0207	0.0206	0.0129	0.0130
	(0.0002)	(0.0002)	(0.0040)	(0.0054)
6. Husband's Education	0.604	0.601	--	--
	(0.0019)	(0.0019)		
7. Parity	-0.121	-0.120	-0.088	-0.062
	(0.000)	(0.000)	(0.010)	(0.015)
State-Mother Cohort Trends	Y	Y	Y	Y
Men Avg. Education		Y		Y

Notes: Standard errors in parentheses. All models include unrestricted mother age effects, unrestricted mother cohort of birth effects, county*year effects (where year is year of birth of the child), median county income in the year when the mother was 17, and percent of the county that is urban in year when the mother was 17. .Models in rows 1 to 6 include all white mothers 24-45 with parity equal to 1. Models in row 7 include white mothers 24 to 45 mothers of any parity. See text for details.

Table 8
Estimates for Mothers 22-45

	OLS (1)	IV (2)	OLS Grouped (3)
SECOND STAGE			
1. Low Birth Weight	-0.0050	-0.0069	-0.0050
	(0.0001)	(0.0034)	(0.0008)
2. Pre-Term Birth	-0.0043	-0.068	-0.0077
	(0.0001)	(0.0039)	(0.0009)
3. Prenatal care	0.0114	0.0283	0.0258
	(0.0001)	(0.0055)	(0.0012)
4. Smoked During Pregnancy	-0.0311	-0.0563	-0.0337
	(0.0004)	(0.0122)	(0.0034)
5. Married	0.0204	0.0057	0.0131
	(0.0002)	(0.0040)	(0.0009)
6. Husband's Education	0.605	--	0.694
	(0.0019)		(0.0081)
7. Parity	-0.121	-0.098	-0.0572
	(0.000)	(0.010)	(0.0030)
FIRST STAGE			
8. Four Year Colleges		0.883	
		(0.037)	
9. Two Year Colleges		0.167	
		(0.035)	

Notes: Standard errors in parentheses. The sample includes all white mothers 22 to 45 with parity equal to 1. See notes to table 5.

Table 9

The Effect of the Transformation of Male-Only Colleges or Female-Only Colleges into Coed Colleges on Maternal Education

Male-Only College to Coed College	0.976
	(0.318)
Female-Only College to Coed College	0.305
	(0.411)

Notes: Standard errors in parentheses. The coefficient in row 1 shows the effect of the transformation of male only four years colleges into coed colleges on mothers' education. There are 106 changes between 1968 and 1983. The coefficient in row 2 shows the effect of the transformation of female-only four years colleges into coed colleges. There are 87 changes between 1968 and 1983. The specification is otherwise similar to the one used in Table 3 (model 1 column 1). The sample includes all white mothers 24 to 45 with parity equal to 1. Data on single sex two year colleges are not available.

Table 10
The Effect of College Openings on Maternal Education
by Number Of Colleges Available in Contiguous Counties

	Coefficient on Four Year Colleges (1)	Coefficient on Two Year Colleges (2)
Model 1: Dividing Counties in Two Groups		
Contiguous counties have less than median number of colleges	1.074	0.299
	(0.086)	(0.069)
Contiguous counties have more than median number of colleges	0.884	0.027
	(0.051)	(0.060)
Model 2: Linear Interaction		
Number of Colleges in own County	1.078	0.288
	(0.076)	(0.062)
Number of Colleges in own County* Number of Colleges in Contiguous Counties	-1.236	-1.154
	(0.465)	(0.498)

Notes: Standard errors in parentheses. The four coefficients for each model come from one regression. All models include main effects for the number of colleges in contiguous counties, unrestricted mother age effects, unrestricted mother cohort of birth effects, county*year effects (where year is year of birth of the child), median county income in year when mother was 17, and percent of the county that is urban in year when the mother was 17. Number of colleges is normalized by cohort size. See text for details.

Table 11
The Percentage of In-State Students, by Age of College

	Four Year Colleges	Two Year Colleges
College is 0-9 years old	91	94
College is 10-19 years old	91	98
College is 20-29 years old	84	95
College is 30-39 years old	79	95
College is 40-49 years old	71	96
College is 50-59 years old	75	94
College is 60-69 years old	75	91
College is 70-79 years old	75	93
College is 80-89 years old	74	94
College is 90-99 years old	77	92
College is 100+ years old	64	94

Notes: Entries show the percentage of students enrolled in four year or two year colleges who are from the state where the college is located. Data are from 1998 IPEDS and include 2303 institutions. We obtain similar results when we look at 1992 or 1994. Data for earlier years are not available.

Table 12
The Effect of College Openings on Enrollment,
by Current Residence and Residence at age 14-16,
1980 and 1990 Census

	Enrolled in School (1)
Model 1	
Four Year Colleges in Location of Residence at age 14-16	0.104
	(0.057)
Two Year Colleges in Location of Residence at age 14-16	0.011
	(0.005)
Model 2	
Four Year Colleges in Current Location of Residence	0.106
	(0.012)
Two Year Colleges in Current Location of Residence	-0.125
	(0.018)

Notes: Standard errors in parentheses. Sample includes white women 19-21 from the 1980 and 1990 5% Censuses of Population. The dependent variable is a dummy equal one if the respondent is enrolled in school. Entries in model 1 show the coefficients on the normalized number of four and two year colleges at age 14-16 in the city where she lived at age 14-16. Entries in model 2 show the coefficients on the normalized number of four and two year colleges when the respondent was 14-16 in the city where she currently lives. Both models include year and location fixed effects. N= 305,225. See text and Data Appendix for details on the samples.

Data Appendix

a) Reporting of Key Variables in the Vital Statistics Data

States began reporting maternal education at different points, so that we do not have a balanced panel of states. Appendix Table A4 shows which states that did not report maternal education over time. In addition, it is important to note that the National Vital Statistics data do not report birth intervals after 1993, so that we cannot incorporate later waves of the Vital Statistics into our panel. Vital Statistics did not begin reporting smoking until 1989. In 1989, 43 states report. States that do not report include CA, IN, LA, NB, NY, OK, SD. By 1992, the states that were not reporting were down to CA, IN, NY, SD. In the most recent years, only CA and NY are missing. Finally, husband education is not reported in detail after 1991. Models examining smoking or husband education are estimated using a smaller data set than those examining other outcomes.

b) Longitudinal Sample

In the longitudinal sample, we include all white women aged 16 to 24 at first birth and 17 to 30 at second birth. We assign all women to a given cell based on the woman's year and state of birth, her county of residence, the first child's birth year, the child's birth month, and the interval between births (which was reported on the birth certificate until 1993). We then drop all cells that have more than one mother. (We have repeated the analysis keeping all cells, and found similar results). Appendix Table A1 shows the distribution of changes in years of schooling. About 22.9% of women in the panel report changes in education between their first and second births. About 11% of women who reported negative changes were deleted from this sample. We obtained similar results when these women were not dropped. Appendix Table A2 shows the percentage of mothers who report an increase in schooling in our Vital Statistics panel and in a comparison sample drawn from the National Longitudinal Survey of Youth (NLSY). To maximize consistency, the NLSY sample includes white mothers who are less than 24 at first birth and who do not change county between births; while the Vital Statistics sample used in this table includes only mothers born between 1957 and 1964.

c) *IV Sample*

The IV sample includes all white women aged 24 to 45 with parity equal 1. Only mothers at first birth are included. While we have no reason to believe that education is systematically misreported in the Vital Statistics data, we have compared education in the Vital Statistics to Census reports. The results are in Appendix Table A3. The first row reports the distribution of education in all years of the Natality Files. These figures correspond to the sample sizes reported at the bottom of Table 1. The low percentage of high-school drop outs and the high percentage of college graduates reflect the fact that the sample includes first-time mothers 24 or older, who tend to be more educated than the average. The remaining rows report the distribution of education in Census years. Columns 1 to 4 show the distribution of maternal education in the Natality Files. Columns 5 to 8 show the distribution of education in the Census. The Natality File figures are generally consistent with Census figures, although the percentage of college graduates tends to be somewhat higher in the Natality Files, while the percentage of high school drops out tend to be lower.

d) Construction of the College Openings Data Set

We began with a listing of accredited two and four year institutions that existed in fall 1996 from the National Center for Education Statistic's Integrated Postsecondary Education Data System (IPEDS) and excluded very small schools (those with less than 200 students in 1996). For four-year colleges, we then searched the Peterson's Guide to Four-Year Colleges (1999) and the Barron's Profiles of American Colleges (1996) for information about the starting date of each college. If the college was listed in both sources and the founding date was the same in the two guides, then we accepted that date. If these conditions were not met, then we searched the internet for information about the founding date of the school.

In reading through college histories, we attempted to choose a date that was as close as possible to the date when actual undergraduate academic instruction began. For example, if the university began as a vocational school and later added academic instruction, we chose the later date. However if the school began as a "Normal school" or a teacher's college we did use that date. Similarly, if the founding date of the school was listed as the date when land was purchased, we used the date when instruction actually began. If the school began as a graduate school or divinity school and later added undergraduate instruction, we used the later date. If a university was formed through the merger of two schools, we used the date at which it opened its doors to the public.

We excluded the following categories of schools: Schools of psychology, law schools, seminaries, Bible colleges and other mainly religious schools, schools that offer only distance or only internet learning, medical schools and medical centers, graduate schools and schools that were purely research facilities, and foreign universities offering degrees in the U.S. in their native languages.

Information on two-year colleges was collected in a similar fashion. There were essentially two types of two-year colleges that existed in 1996: Those that offered one and/or two year degree programs as well as transfer programs to four year schools, and those that offered only very specific vocational programs. However, we found that most two-year colleges offering broader programs now began as vocational institutions. Thus, for the sake of consistency, we decided to keep all two year colleges, excluding only hairdressing and beauty schools. Excluding institutions that did not offer an Associates Degree had very little effect on our results, however.

An important potential problem is that some schools that existed over part of our sample period may have exited the sample by 1996. In order to assess the severity of this problem, we used data generously supplied by David Card about colleges that were in existence at 5 year intervals from 1960 on. This information was compiled using the Department of Education's CASPAR data base. We estimate that less than three percent of schools that were in existence at some point had exited by 1996. Adding these schools to our data base, had virtually no effect on the estimates reported in the paper both because the number of schools that closed is small, and because schools that closed were generally very small institutions.

Our instrumental variable is the number of colleges which existed in the woman's county when she was 17 years old divided by the estimated number of 18 to 22 year olds in the county in that year. Since population numbers are not available by age group, county, and year, we impute the number of 18 to 22 year olds in the county using the county population in each year (interpolating between Census years) and information about the number of 18 to 22 year olds in the state in each Census year. That is, we assume that 18 to 22 year olds are distributed across each state roughly in proportion to county populations. We have experimented with two

alternative imputation strategies, and found similar results.

In order to identify conversions of single-sex to coed colleges, we started with the 1968 to 1983 waves of the HEGIS (Higher Education General Information Survey) data available from the National Center for Education Statistics. This data identifies each school as all male, all female, coed, or "coordinate" (which indicates that the institution had integrated classes even if it did not formally admit both men and women). We matched this data to the list of four year colleges that we had already constructed from the 1996 IPEDS data. This enabled us to identify which of the colleges began as single-sex institutions, and the year that they went coed, if the year was between 1969 and 1983. For colleges that were still single sex in 1983, but had become coed by 1996, we checked college web sites in order to determine the year of conversion. We found that the typical school that changed from single-sex to coed was a small religious institution (rather than a Princeton or a Dartmouth). Also, we did not find any instances of two-year colleges that began as single-sex institutions. Colleges that were listed as "coordinate" at some point, generally held this status for one or two years before becoming coed. We used the first date that they were listed as coed as the date of conversion.

Reliance on the HEGIS data means that we may have missed conversions of single-sex colleges that occurred before 1968. Harwarth et al. (2002) cite data from the Education Directories produced by the U.S. Dept. of Education which show that there were 276 women's colleges in 1945, 248 women's colleges in 1955, and 252 women's colleges in 1960. They also cite problems with these data which suggest that one should not place a great deal of weight on small variations in the number of colleges reporting. The HEGIS data indicate that there were 258 women's colleges in 1969, but that the number declined rapidly after that (to 230 in 1969, 194 in 1970 and to 103 by 1983). Thus, our procedures capture the era of greatest change in the number of female colleges after 1955, though we miss one earlier period with many conversions between 1945 and 1955.

e) *Census Data*

We use the 1980 and 1990 5% sample of the Census (PUMS) to identify place residence in 1975 and 1980 (in 1980 Census) and 1985 and 1990 (in 1990 Census). The main limitation of the data is that counties are not reported. The smallest geographical identifiers available in the PUMS are called Public Use Microdata Areas, or PUMAs and identify areas of no less than 100,000 population. In rural areas, PUMAs are larger than counties, while in densely populated areas, PUMAs are smaller than counties. For example, for the nine-county San Francisco Bay Area, there are 48 PUMAs. The PUMS also report metropolitan areas identifiers, although the definition changes between 1980 and 1990. We assign individuals located in metropolitan areas the total number of colleges and of individuals 18-22 in the metropolitan area. We make the metropolitan area definition in 1990 consistent with the one in 1980. The codes are available on request. We assign individuals located outside metropolitan areas, the total number of colleges and of individuals 18-22 in the PUMA. In the1980 sample, not everyone was asked about location of residence in 1975, so that the sample size is smaller in 1980 than in 1990.

Appendix Table A1: Distribution of Changes in Years of Schooling for Mothers in the Panel

	Percent of Sample	Percent HS Drop Out at Second Birth	Percent HS Graduate at Second Birth	Percent with Some College at Second Birth	Percent College Grad. at Second Birth
Change in years of schooling = 0	77.1	16	62	14	6
Change in years of schooling = 1	14.5	22	33	37	6
Change in years of schooling = 2	5.9	9	44	37	8
Change in years of schooling = 3	1.6	4	43	36	15
Change in years of schooling = 4	0.6	1	23	19	56
Change in years of schooling = 5	0.09	1	13	30	54
Change in years of schooling = 6	0.01	0	18	34	46

Notes: The sample includes mothers who were 16-24 at their first birth and 17-30 at their second birth. The sample was obtained as follows: first, women with parity =1 were selected. Then, women with parity = 2 were matched to women with parity =1 based on county, date of birth of first child (month and year), mother's age at first birth, mother's year of birth and mother's state of birth. Finally, only cells with size equal 1 in both years were kept (one-to-one matches). The sample is therefore an actual panel of mothers with two children, where the same mother is observed at the time of the first birth and the time of the second birth. See text for details. Observations for which the change in schooling is negative are dropped.

Appendix Table A2: Comparison of Changes in Schooling in Vital Statistics and NLSY

	Percentage of Mothers who Report Positive Increases in Schooling Between Births, by Age at First Birth			
	All	First Birth <= 18	First Birth 19-20	First Birth 21-24
Vital Statistics Panel	0.21	0.38	0.17	0.16
NLSY	0.17	0.36	0.15	0.07

Note: For consistency with the Vital Statistics sample, the NLSY sample includes white mothers who are less than 24 at first birth and who do not change county between births. The Vital Statistics sample used in this table is similar to the one used in Tables 2 and A2, but it includes only mothers born between 1957 and 1964 to maximize consistency with the NLSY. The sample sizes in columns 1 to 4 for the NLSY sample are, respectively, 691, 177, 201 and 313.

Appendix Table A3

Comparison of Education of Mothers of any Parity in the Natality Files and the Census

	NATALITY FILES				CENSUS			
	College	Some College	High School	Less than High School	College	Some College	High School	Less than High School
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Years (1970-99)	40%	27%	30%	3%				
1970, 1980, 1990	38%	26%	32%	4%	33%	30%	31%	5%
1990	40%	27%	30%	3%	35%	34%	26%	4%
1980	37%	26%	34%	3%	30%	26%	36%	6%
1970	33%	20%	41%	6%	26%	18%	42%	12%

Notes: Columns 1 to 4 show the distribution of maternal education in the Natality Files. First time white mothers 24-45 are included. Columns 5 to 8 show the distribution of education in the Census. First time white mothers age 24-45 who had a child in the last year are included.

Appendix Table A4: States Missing Maternal Education in the Vital Statistics Data, By Year

1969	AL	AR	CA	CT	DE	DC	FL	GA	ID	MD	NM	PA	TX	WA
1970	AL	AR	CA	CT	DE	DC		GA	ID	MD	NM	PA	TX	WA
1971	AL	AR	CA	CT	DE	DC		GA	ID	MD	NM	PA	TX	WA
1972	AL	AR	CA	CT		DC		GA	ID	MD	NM	PA	TX	WA
1973	AL	AR	CA						ID	MD	NM	PA	TX	WA
1974	AL	AR	CA						ID		NM	PA	TX	WA
1975	AL	AR	CA						ID		NM	PA	TX	WA
1976		AR	CA						ID		NM		TX	WA
1977		AR	CA						ID		NM		TX	WA
1978											NM		TX	WA
1979											NM		TX	WA
1980-1988													TX	WA
1989-1991														WA
1992+						All states report.								

Appendix 2: IV and OLS in a Model with Spillovers

We show that OLS and IV estimates differ when spillovers are introduced. This is a simplified version of an argument that originally appeared in Lochner and Moretti (2002). Consider the a simple model of birth outcomes with county-level spillovers:

$$y_{ic} = \alpha + \beta S_{ic} + \gamma S_c + \varepsilon_{ic},$$

where y_{ic} is the birth outcome of mother i in county c; S_{ic} is a dummy equal one if individual i is a college graduate; S_c represents college graduation rate in county c; and ε_{is} is a mean zero random error term assumed to be independent of S_{is} and S_s. To simplify the analysis, assume we have a balanced panel with n individuals in each of C Counties. We also define S the average graduation rate in the entire economy.

OLS estimates of a regression of y_{ic} on S_{ic} is

$$\begin{aligned}
\hat{\beta}_{ols} &= \beta + \gamma\left(\frac{\sum_c\sum_i(S_{ic}-S)(S_c-S)}{\sum_c\sum_i(S_{ic}-S)^2}\right) + \frac{\sum_c\sum_i \epsilon_{ic}^2}{\sum_c\sum_i(S_{ic}-S)^2} \\
&= \beta + \gamma\left(\frac{\sum_c(S_c-S)(S_c-S)}{\sum_c\sum_i(S_{ic}-S)^2}\right) + \frac{\sum_c\sum_i \epsilon_{ic}^2}{\sum_c\sum_i(S_{ic}-S)^2} \\
&\to \beta + \gamma\left(\frac{V(S_c)}{V(S_{ic})}\right).
\end{aligned}$$

IV estimates, using a valid county-level instrument z_c (with sample mean z) are:

$$\begin{aligned}
\hat{\beta}_{IV} &= \beta + \gamma\left(\frac{\sum_c\sum_i(z_c-z)(S_c-S)}{\sum_c\sum_i(z_c-z)(S_{ic}-S)}\right) + \frac{\sum_c\sum_i(z_c-z)\epsilon_{ic}}{\sum_c\sum_i(z_c-z)(S_{ic}-S)} \\
&= \beta + \gamma\left(\frac{\sum_c(z_c-z)(S_c-S)}{\sum_c(z_c-z)(S_c-S)}\right) + \frac{\sum_c\sum_i(z_c-z)\epsilon_{ic}}{\sum_c\sum_i(z_c-z)(S_{ic}-S)} \\
&\to \beta + \gamma.
\end{aligned}$$

For $\gamma > 0$, we will observe $\beta_{IV} \geq \beta_{ols}$ since $V(d_s) \leq V(d_{is})$. For small cross-county variation in graduation rates (i.e. $V(d_s) \approx 0$), OLS will estimate the own-effect of college on birth outcomes, while IV estimates using county-level instruments will estimate the combined own-effect and spillover effect of college.

Defining y_c the average birth outcome in a county, estimates from a regression of y_c on average graduation rates in a county will produce estimates of the combined own and spillover effects, since $y_c = \alpha + (\beta + \gamma)S_c + \varepsilon_c$. (Note that for simplicity, here we assume that the instrument varies at the county level, while in fact the instrument used in this paper varies at the county and cohort of birth level. Obviously the same argument applies when c indexes county×cohort of birth)

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

UD035607

I. DOCUMENT IDENTIFICATION:

Title: Mothers Education and the Intergenerational Transmission of Human Capital	
Author(s): Janet Currie, Enrico Moretti.	
Corporate Source: National Bureau of Economic Research	Publication Date: Dec. 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☑ Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐ Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐ Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: [signature]	Printed Name/Position/Title: Professor Janet Currie	
Organization/Address: Dept. of Economics UCLA, 405 Hilgard Ave Los Angeles CA 90095 1477	Telephone: 310 2068380	FAX: 3104716862
	E-Mail Address: currie@ssr.ucla.edu	Date: Apr 3/03

(over)